

日本国特許庁

JAPAN PATENT OFFICE

PCT/JP03/05419

28.04.03

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office

出願年月日

Date of Application:

2002年 9月12日

出願番号

Application Number:

特願2002-267436

[ST.10/C]:

[JP2002-267436]

出願人

Applicant(s):

松下電器産業株式会社

REC'D 20 JUN 2003

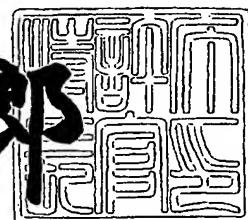
WIPO PCT

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

2003年 6月 2日

特許庁長官
Commissioner,
Japan Patent Office

太田信一郎



BEST AVAILABLE COPY

出証番号 出証特2003-3042104

【書類名】 特許願
【整理番号】 2900645205
【提出日】 平成14年 9月12日
【あて先】 特許庁長官殿
【国際特許分類】 H04B 7/26
【発明者】
【住所又は居所】 神奈川県横浜市港北区綱島東四丁目3番1号 松下通信
工業株式会社内
【氏名】 押切 正浩
【特許出願人】
【識別番号】 000005821
【氏名又は名称】 松下電器産業株式会社
【代理人】
【識別番号】 100105050
【弁理士】
【氏名又は名称】 鶴田 公一
【手数料の表示】
【予納台帳番号】 041243
【納付金額】 21,000円
【提出物件の目録】
【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1
【包括委任状番号】 9700376
【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音響符号化装置及び音響符号化方法

【特許請求の範囲】

【請求項1】 入力信号のサンプリングレートを下げるダウンサンプリング手段と、サンプリングレートが下げられた入力信号を符号化する基本レイヤ符号化手段と、符号化された入力信号を復号化して復号信号を得る復号化手段と、前記復号信号のサンプリングレートを入力時の入力信号のサンプリングレートと同一のレートに上げるアップサンプリング手段と、入力時の入力信号とサンプリングレートが上げられた復号信号との差分から誤差信号を得る減算手段と、サンプリングレートが上げられた復号信号に基づいて前記誤差信号を符号化する対象の周波数を決定する周波数決定手段と、前記周波数にある前記差分信号を符号化する拡張レイヤ符号化手段と、を具備することを特徴とする音響符号化装置。

【請求項2】 前記基本レイヤ符号化手段は、符号励振線形予測法を用いて入力信号を符号化することを特徴とする請求項1に記載の音響符号化装置。

【請求項3】 前記拡張レイヤ符号化手段は、前記差分信号を時間領域から周波数領域に直交変換し、変換後の前記差分信号を符号化することを特徴とする請求項1または請求項2に記載の音響符号化装置。

【請求項4】 聴覚に寄与しない振幅値を表す聴覚マスキングを算出する聴覚マスキング手段を具備し、前記拡張レイヤ符号化手段は、前記周波数決定手段において前記聴覚マスキング内の信号を符号化の対象としないように符号化する対象を決定して前記誤差信号のスペクトルである誤差スペクトルを符号化することを特徴とする請求項1から請求項3のいずれかに記載の音響符号化装置。

【請求項5】 前記聴覚マスキング手段は、サンプリングレートが上げられた復号信号を周波数領域の係数に変換する周波数領域変換手段と、前記周波数領域の係数を用いて推定聴覚マスキングを算出する推定聴覚マスキング算出手段と、前記復号信号のスペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求める決定手段と、を具備し、前記拡張レイヤ符号化手段は、前記周波数に位置する前記誤差スペクトルを符号化することを特徴とする請求項4に記載の音響符号化装置。

【請求項6】 前記聴覚マスキング手段は、前記周波数領域の係数を用いて推定誤差スペクトルを算出する推定誤差スペクトル算出手段を具備し、前記決定手段は、前記推定誤差スペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求ることを特徴とする請求項5に記載の音響符号化装置。

【請求項7】 前記聴覚マスキング手段は、前記推定聴覚マスキング算出手段において算出された推定聴覚マスキングを平滑化する修正手段を具備し、前記決定手段は、前記復号信号のスペクトルまたは前記推定誤差スペクトルの振幅値が平滑化された前記推定聴覚マスキングの振幅値を超える周波数を求ることを特徴とする請求項5または請求項6に記載の音響符号化装置。

【請求項8】 前記拡張レイヤ符号化手段は、推定誤差スペクトルまたは誤差スペクトルのいずれかと聴覚マスキングまたは推定聴覚マスキングのいずれかとの振幅値の差を周波数毎に算出し、前記振幅値の差の大きさに基づいて符号化の情報量を決定することを特徴とする請求項5から請求項7のいずれかに記載の音響符号化装置。

【請求項9】 前記拡張レイヤ符号化手段は、前記決定手段で求めた周波数に加えて、予め定められた帯域にある前記誤差スペクトルを符号化することを特徴とする請求項4から請求項8のいずれかに記載の音響符号化装置。

【請求項10】 符号化側において入力信号を所定の基本フレーム単位で符号化した第1符号化コードを復号化して第1復号信号を得る基本レイヤ復号化手段と、第1復号信号のサンプリングレートを第2復号信号のサンプリングレートと同一のサンプリングレートに上げるアップサンプリング手段と、前記アップサンプリングされた第1復号信号に基づいて入力信号と符号化側において第1符号化コードを復号した信号との残差信号を符号化した第2符号化コードを復号化する対象の周波数を決定する周波数決定手段と、前記周波数の情報を用いて前記第2符号化コードを復号化して第2復号信号を得る拡張レイヤ復号化手段と、前記第2復号信号とサンプリングレートを上げられた第1復号信号を加算する加算手段と、を具備することを特徴とする音響復号化装置。

【請求項11】 基本レイヤ復号化手段は、符号励振線形予測法を用いて第1符号化コードを復号化することを特徴とする請求項10に記載の音響復号化裝

置。

【請求項12】 拡張レイヤ復号化手段は、第2符号化コードを復号化した信号を周波数領域から時間領域に直交変換することを特徴とする請求項10または請求項11に記載の音響復号化装置。

【請求項13】 聴覚に寄与しない振幅値を表す聴覚マスキングを算出する聴覚マスキング手段を具備し、前記拡張レイヤ復号化手段は、前記周波数決定手段において前記聴覚マスキング内の信号を復号化の対象としないように復号化する対象を決定することを特徴とする請求項10から請求項12のいずれかに記載の音響復号化装置。

【請求項14】 前記聴覚マスキング手段は、サンプリングレートの上げられた基本レイヤの復号信号を周波数領域の係数に変換する周波数領域変換手段と、前記周波数領域の係数を用いて推定聴覚マスキングを算出する推定聴覚マスキング算出手段と、前記復号信号のスペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求める決定手段と、を具備し、前記拡張レイヤ復号化手段は、前記周波数に位置する前記誤差スペクトルを復号化することを特徴とする請求項13に記載の音響復号化装置。

【請求項15】 前記聴覚マスキング手段は、前記周波数領域の係数を用いて推定誤差スペクトルを算出する推定誤差スペクトル算出手段を具備し、前記決定手段は、前記推定誤差スペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求ることを特徴とする請求項14に記載の音響復号化装置。

【請求項16】 前記聴覚マスキング手段は、前記推定聴覚マスキング算出手段において算出された推定聴覚マスキングを平滑化する修正手段を具備し、前記決定手段は、前記復号信号のスペクトルまたは前記推定誤差スペクトルの振幅値が平滑化された前記推定聴覚マスキングの振幅値を超える周波数を求ることを特徴とする請求項14または請求項15に記載の音響復号化装置。

【請求項17】 前記拡張レイヤ復号化手段は、推定誤差スペクトルまたは誤差スペクトルのいずれかと聴覚マスキングまたは推定聴覚マスキングのいずれかとの振幅値の差を周波数毎に算出し、前記振幅値の差の大きさに基づいて復号化の情報量を決定することを特徴とする請求項14から請求項16のいずれかに

記載の音響復号化装置。

【請求項18】 前記拡張レイヤ復号化手段は、前記決定手段で求めた周波数に加えて、予め定められた帯域にある前記誤差スペクトルとを復号化することを特徴とする請求項13から請求項17のいずれかに記載の音響復号化装置。

【請求項19】 音響信号を電気的信号に変換する音響入力手段と、この音響入力手段から出力された信号をデジタル信号に変換するA/D変換手段と、このA/D変換手段から出力されたデジタル信号を符号化する請求項1から請求項9のいずれかに記載の音響符号化装置と、この符号化装置から出力された符号化コードを無線周波数の信号に変調するRF変調手段と、このRF変調手段から出力された信号を電波に変換して送信する送信アンテナとを具備することを特徴とする音響信号送信装置。

【請求項20】 電波を受信する受信アンテナと、この受信アンテナに受信された信号を復調するRF復調手段と、このRF復調手段にて得られた情報を復号する請求項10から請求項18のいずれかに記載の音響復号化装置と、この復号化装置から出力された信号をアナログ信号に変換するD/A変換手段と、このD/A変換手段から出力された電気的信号を音響信号に変換する音響出力手段とを具備することを特徴とする音響信号受信装置。

【請求項21】 請求項19記載の音響信号送信装置あるいは請求項20記載の音響信号受信装置の少なくとも一方を具備することを特徴とする通信端末装置。

【請求項22】 請求項19記載の音響信号送信装置あるいは請求項20記載の音響信号受信装置の少なくとも一方を具備することを特徴とする基地局装置。

【請求項23】 本発明の音響符号化方法は、符号化側において、サンプリングレートが下げられた入力信号を符号化して第1符号化コードを作成し、前記第1符号化コードを復号化した第1復号信号のサンプリングレートを入力時の入力信号のサンプリングレートと同一のレートに上げ、サンプリングレートが上げられた復号信号に基づいて前記誤差信号を符号化する対象の周波数を決定し、入力時の入力信号とサンプリングレートが上げられた復号信号との差分信号のうち

前記周波数にある前記差分信号を符号化して第2符号化コードを作成し、復号化側において、前記第1符号化コードを復号化して第2復号信号を得て、前記第2復号信号のサンプリングレートを第3復号信号のサンプリングレートと同一のレートに上げ、このサンプリングレートの上げられた第2復号信号に基づいて第2符号化コードを復号化する対象の周波数を決定し、前記周波数の情報を用いて前記第2符号化コードを復号化して第3復号信号を得て、前記サンプリングレートの上げられた第2復号信号と前記第3復号信号とを加算することを特徴とする音響符号化方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、楽音信号または音声信号などの音響信号を高能率に圧縮符号化する音響符号化装置及び音響符号化方法に関し、特に符号化コードの一部からでも楽音や音声を復号することができるスケーラブル符号化を行う音響符号化装置及び音響符号化方法に関する。

【0002】

【従来の技術】

楽音信号または音声信号を低ビットレートで圧縮する音響符号化技術は、移動体通信における電波等の伝送路容量及び記録媒体の有効利用のために重要である。音声信号を符号化する音声符号化に、ITU(International Telecommunication Union)で規格化されているG726、G729などの方式がある。これらの方は、狭帯域信号(300Hz～3.4kHz)を対象とし、8kbit/s～32kbit/sのビットレートで高品質に符号化できる。

【0003】

また、広帯域信号(50Hz～7kHz)を符号化する標準方式としてITUのG722、G722.1や、3GPP(The 3rd Generation Partnership Project)のAMR-WBなどが存在する。これらの方は、6.6kbit/s～6.4kbit/sのビットレートで広帯域音声信号を高品質に符号化できる。

【0004】

音声信号を低ビットレートで高能率に符号化を行う有効な方法に、C E L P (Code Excited Linear Prediction)がある。C E L Pは、人間の音声生成モデルを工学的に模擬したモデルに基づき、乱数やパルス列で表される励振信号を周期性の強さに対応するピッチフィルタと声道特性に対応する合成フィルタに通し、その出力信号と入力信号の二乗誤差が聴覚特性の重み付けの下で最小になるよう符号化パラメータを決定する方法である（例えば、非特許文献1参照）。

【0005】

最近の標準音声符号化方式の多くがC E L Pに基づいており、例えばG 729は、8 k b i t / s のビットレートで狭帯域信号の符号化でき、A M R - W B は6. 6 k b i t / s ~ 23. 85 k b i t / s のビットレートで広帯域信号を符号化できる。

【0006】

一方で、楽音信号を符号化する楽音符号化の場合、M P E G (Moving Picture Expert Group)で規格化されているレイヤ3方式やA A C方式のように、楽音信号を周波数領域に変換し、聴覚心理モデルを利用して符号化を行う変換符号化が一般的である。これらの方程式は、サンプリングレートが44. 1 k H z の信号に対しチャネル当たり64 k b i t / s ~ 96 k b i t / s のビットレートでほとんど劣化が生じないことが知られている。

【0007】

しかしながら、音声信号が主体で、背景に音楽や環境音が重畠している信号を符号化する場合、音声符号化方式を適用すると背景部の音楽や環境音の影響で、背景部の信号のみならず音声信号も劣化してしまい全体的な品質が低下するという問題がある。これは、音声符号化方式が、C E L Pという音声モデルに特化した方式を基本にしているために生じる問題である。また、音声符号化方式が対応できる信号帯域は高々7 k H zまでであり、それ以上の高域を持つ信号に対しては構成上十分に対応しきれないという問題がある。

【0008】

一方で、楽音符号化は、音楽に対して高品質に符号化を行うことができるので、前述したような背景に音楽や環境音がある音声信号についても十分な品質を得

ることができる。また、楽音符号化は、対象となる信号の帯域もCD品質であるサンプリングレートが22kHz程度の信号まで対応可能である。

【0009】

その反面、高品質な符号化を実現するためにはビットレートを高くして使用する必要があり、仮にビットレートを32kbit/s程度まで低く抑えると復号信号の品質が大きく低下するという問題がある。そのため、伝送レートの低い通信網で使用できないという問題がある。

【0010】

上述した問題を回避するためにこれらの技術を組み合わせて、最初に入力信号を基本レイヤにてCELPで符号化し、次にその復号信号を入力信号から減算して得られる誤差信号を求め、この信号を拡張レイヤにて変換符号化を行うスケーラブル符号化が考えられる。

【0011】

この方法では、基本レイヤはCELPを用いているため音声信号を高品質に符号化でき、かつ拡張レイヤは基本レイヤで表しきれない背景の音楽や環境音、基本レイヤでカバーする周波数帯よりも高い周波数成分の信号を効率よく符号化することができる。さらにこの構成によればビットレートを低く抑えることができる。加えて、この構成によれば、符号化コードの一部つまり基本レイヤの符号化コードのみから音響信号を復号することが可能であり、このようなスケーラブル機能は伝送容量の異なる複数のネットワークに対するマルチキャストの実現に有効である。

【0012】

しかしながら、音声ではなく音楽を入力したときに十分な品質を確保するためには、拡張レイヤへのビット配分を多くする必要があり、その結果ビットレートが高くなってしまうという問題がある。

【0013】

基本レイヤの符号化において、CELPのような音声に特化した符号化方式を用いているが、このCELPでは音楽に対する符号化効率が高くなかった。音楽信号を符号化すると、入力信号と基本レイヤの復号信号との誤差信号（拡張レイヤの

入力信号) のパワーが大きくなるので、パワーの大きくなった誤差信号に対応するため拡張レイヤに多くのビットを配分して、最終的な復号信号の品質を上げる必要があった。

【0014】

この問題を解決するために、拡張レイヤで聴覚マスキングを利用して符号化効率を上げることが考えられる。聴覚マスキングとは、ある信号が与えられたときその信号の周波数の近傍に位置する信号が聞こえなくなる(マスクされる)という人間の聴覚特性を利用したものである。

【0015】

図28は、音響(音楽)信号のスペクトルの一例を示す図である。図28において、実線は聴覚マスキングを表し、破線は誤差スペクトルを表している。ここでいう誤差スペクトルとは、入力信号と基本レイヤの復号信号との誤差信号(拡張レイヤの入力信号)のスペクトルを指す。

【0016】

図28の斜線部で表される誤差スペクトルは、聴覚マスキングよりも振幅値が小さいため人間の聴覚では聞こえず、それ以外の領域では誤差スペクトルの振幅値が聴覚マスキングを超えてるので量子化歪が知覚される。

【0017】

そこで、拡張レイヤでは図28の白地部に含まれる誤差スペクトルを符号化してその領域の量子化歪が聴覚マスキングよりも小さくなるようにすればよい。また、斜線部に属する係数は既に聴覚マスキングよりも小さくので量子化する必要がない。

【0018】

【非特許文献1】

"Code-Excited Linear Prediction (CELP): high quality speech at very low bit rates", Proc. ICASSP 85, pp.937-940, 1985.

【0019】

【発明が解決しようとする課題】

しかしながら、従来の装置においては、聴覚マスキングにより量子化が必要な

周波数の情報を伝送する必要があり、伝送する情報量が増加してビットレートを低くすることができないという問題がある。

【0020】

本発明はかかる点に鑑みてなされたものであり、声が主体で背景に音楽や雑音が重畠しているような信号であっても、低ビットレートで高品質に符号化を行うことができる音響符号化装置及び音響符号化方法を提供することを目的とする。

【0021】

【課題を解決するための手段】

本発明の音響符号化装置は、入力信号のサンプリングレートを下げるダウンサンプリング手段と、サンプリングレートが下げられた入力信号を符号化する基本レイヤ符号化手段と、符号化された入力信号を復号化して復号信号を得る復号化手段と、前記復号信号のサンプリングレートを入力時の入力信号のサンプリングレートと同一のレートに上げるアップサンプリング手段と、入力時の入力信号とサンプリングレートが上げられた復号信号との差分から誤差信号を得る減算手段と、サンプリングレートが上げられた復号信号に基づいて前記誤差信号を符号化する対象の周波数を決定する周波数決定手段と、前記周波数にある前記差分信号を符号化する拡張レイヤ符号化手段と、を具備する構成を探る。

【0022】

この構成によれば、符号化信号の復号化した信号から拡張レイヤの符号化の対象となる周波数を決定することにより、符号化側から復号化側に伝送する基本レイヤの符号化信号のみで拡張レイヤの符号化の対象となる周波数を決定することができ、符号化側から復号化側にこの周波数の情報を伝送する必要がなくなり、低ビットレートで高品質に符号化を行うことができる。

【0023】

本発明の音響符号化装置は、前記基本レイヤ符号化手段は、符号励振線形予測法を用いて入力信号を符号化する構成を探る。

【0024】

この構成によれば、送信側において、基本レイヤにC E L Pを適用して入力信号を符号化し、受信側において、この符号化した入力信号にC E L Pを適用して

復号することにより、低ビットレートで高品質な基本レイヤを実現することができる。

【0025】

本発明の音響符号化装置は、前記拡張レイヤ符号化手段は、前記差分信号を時間領域から周波数領域に直交変換し、変換後の前記差分信号を符号化する構成を探る。

【0026】

この構成によれば、差分信号を時間領域から周波数領域に変換し、変換後の信号について基本レイヤの符号化によりカバーできない周波数領域を拡張レイヤで符号化することにより、音楽のようにスペクトルの変化が大きい信号にも対応することができる。

【0027】

本発明の音響符号化装置は、聴覚に寄与しない振幅値を表す聴覚マスキングを算出する聴覚マスキング手段を具備し、前記拡張レイヤ符号化手段は、前記周波数決定手段において前記聴覚マスキング内の信号を符号化の対象としないように符号化する対象を決定して前記誤差信号のスペクトルである誤差スペクトルを符号化する構成を探る。

【0028】

本発明の音響符号化装置は、前記聴覚マスキング手段は、サンプリングレートが上げられた復号信号を周波数領域の係数に変換する周波数領域変換手段と、前記周波数領域の係数を用いて推定聴覚マスキングを算出する推定聴覚マスキング算出手段と、前記復号信号のスペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求める決定手段と、を具備し、前記拡張レイヤ符号化手段は、前記周波数に位置する前記誤差スペクトルを符号化する構成を探る。

【0029】

これらの構成によれば、マスキング効果の特性を利用して、入力信号のスペクトルから聴覚マスキングを算出し、拡張レイヤの符号化において、量子化歪をこのマスキング値以下になるように量子化を行うことにより、品質の劣化を伴わずに量子化の対象となるMDCT係数の数を減らすことができ、低ビットレートで

高品質に符号化を行うことができる。

【0030】

本発明の音響符号化装置は、前記聴覚マスキング手段は、前記周波数領域の係数を用いて推定誤差スペクトルを算出する推定誤差スペクトル算出手段を具備し、前記決定手段は、前記推定誤差スペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求める構成を探る。

【0031】

この構成によれば、基本レイヤの復号信号のスペクトルから推定した残差スペクトルを平滑化することにより、推定誤差スペクトルを残差スペクトルに近似することができ、拡張レイヤにて誤差スペクトルを効率よく符号化することができる。

【0032】

本発明の音響符号化装置は、前記聴覚マスキング手段は、前記推定聴覚マスキング算出手段において算出された推定聴覚マスキングを平滑化する修正手段を具備し、前記決定手段は、前記復号信号のスペクトルまたは前記推定誤差スペクトルの振幅値が平滑化された前記推定聴覚マスキングの振幅値を超える周波数を求める構成を探る。

【0033】

この構成によれば、基本レイヤ復号信号の振幅スペクトルから推定した推定聴覚マスキングを、基本レイヤ符号化器の符号化コードの情報を基に修正を加えることにより、推定聴覚マスキングの精度を向上させることができ、結果拡張レイヤにて誤差スペクトルを効率よく符号化することができる。

【0034】

本発明の音響符号化装置は、前記拡張レイヤ符号化手段は、推定誤差スペクトルまたは誤差スペクトルのいずれかと聴覚マスキングまたは推定聴覚マスキングのいずれかとの振幅値の差を周波数毎に算出し、前記振幅値の差の大きさに基づいて符号化の情報量を決定する構成を探る。

【0035】

この構成によれば、拡張レイヤでの符号化において、推定誤差スペクトルが推

定聴覚マスキングを超える量が大きい周波数に多くの情報量を配分して符号化することにより、量子化効率の向上を図ることができる。

【0036】

本発明の音響符号化装置は、前記拡張レイヤ符号化手段は、前記決定手段で求めた周波数に加えて、予め定められた帯域にある前記誤差スペクトルを符号化する構成を探る。

【0037】

この構成によれば、符号化の対象として選択されにくいが聴覚的に重要な帯域を強制的に量子化することにより、本来符号化の対象として選択されるべき周波数が選択されない場合でも、聴覚的に重要な帯域に含まれる周波数に位置する誤差スペクトルは必ず量子化されることになり、品質の改善することができる。

【0038】

本発明の音響復号化装置は、符号化側において入力信号を所定の基本フレーム単位で符号化した第1符号化コードを復号化して第1復号信号を得る基本レイヤ復号化手段と、第1復号信号のサンプリングレートを第2復号信号のサンプリングレートと同一のサンプリングレートに上げるアップサンプリング手段と、前記アップサンプリングされた第1復号信号に基づいて入力信号と符号化側において第1符号化コードを復号した信号との残差信号を符号化した第2符号化コードを復号化する対象の周波数を決定する周波数決定手段と、前記周波数の情報を用いて前記第2符号化コードを復号化して第2復号信号を得る拡張レイヤ復号化手段と、前記第2復号信号とサンプリングレートを上げられた第1復号信号を加算する加算手段と、を具備する構成を探る。

【0039】

この構成によれば、基本レイヤの符号化信号の復号化した信号から拡張レイヤの符号化の対象となる周波数を決定することにより、符号化側から復号化側に伝送する基本レイヤの符号化信号のみで拡張レイヤの符号化の対象となる周波数を決定することができ、符号化側から復号化側にこの周波数の情報を伝送する必要がなくなり、低ビットレートで高品質に符号化を行うことができる。

【0040】

本発明の音響復号化装置は、基本レイヤ復号化手段は、符号励振線形予測法を用いて第1符号化コードを復号化する構成を探る。

【0041】

この構成によれば、送信側において、基本レイヤにCELPを適用して入力信号を符号化し、受信側において、この符号化した入力信号にCELPを適用して復号することにより、低ビットレートで高品質な基本レイヤを実現することができる。

【0042】

本発明の音響復号化装置は、拡張レイヤ復号化手段は、第2符号化コードを復号化した信号を周波数領域から時間領域に直交変換する構成を探る。

【0043】

この構成によれば、差分信号を時間領域から周波数領域に変換し、変換後の信号について基本レイヤの符号化によりカバーできない周波数領域を拡張レイヤで符号化することにより、音楽のようにスペクトルの変化が大きい信号にも対応することができる。

【0044】

本発明の音響復号化装置は、聴覚に寄与しない振幅値を表す聴覚マスキングを算出する聴覚マスキング手段を具備し、前記拡張レイヤ復号化手段は、前記周波数決定手段において前記聴覚マスキング内の信号を復号化の対象としないように復号化する対象を決定する構成を探る。

【0045】

本発明の音響復号化装置は、前記聴覚マスキング手段は、サンプリングレートの上げられた基本レイヤの復号信号を周波数領域の係数に変換する周波数領域変換手段と、前記周波数領域の係数を用いて推定聴覚マスキングを算出する推定聴覚マスキング算出手段と、前記復号信号のスペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求める決定手段と、を具備し、前記拡張レイヤ復号化手段は、前記周波数に位置する前記誤差スペクトルを復号化する構成を探る。

【0046】

これらの構成によれば、マスキング効果の特性を利用して、入力信号のスペクトルから聴覚マスキングを算出し、拡張レイヤの符号化において、量子化歪をこのマスキング値以下になるように量子化を行うことにより、品質の劣化を伴わずに量子化の対象となるMDCT係数の数を減らすことができ、低ビットレートで高品質に符号化を行うことができる。

【0047】

本発明の音響復号化装置は、前記聴覚マスキング手段は、前記周波数領域の係数を用いて推定誤差スペクトルを算出する推定誤差スペクトル算出手段を具備し、前記決定手段は、前記推定誤差スペクトルの振幅値が前記推定聴覚マスキングの振幅値を超える周波数を求める構成を探る。

【0048】

この構成れば、基本レイヤの復号信号のスペクトルから推定した残差スペクトルを平滑化することにより、推定誤差スペクトルを残差スペクトルに近似することができ、拡張レイヤにて誤差スペクトルを効率よく符号化することができる。

【0049】

本発明の音響復号化装置は、前記聴覚マスキング手段は、前記推定聴覚マスキング算出手段において算出された推定聴覚マスキングを平滑化する修正手段を具備し、前記決定手段は、前記復号信号のスペクトルまたは前記推定誤差スペクトルの振幅値が平滑化された前記推定聴覚マスキングの振幅値を超える周波数を求める構成を探る。

【0050】

この構成によれば、基本レイヤ復号信号の振幅スペクトルから推定した推定聴覚マスキングを、基本レイヤ符号化器の符号化コードの情報を基に修正を加えることにより、推定聴覚マスキングの精度を向上させることができ、結果拡張レイヤにて誤差スペクトルを効率よく符号化することができる。

【0051】

本発明の音響復号化装置は、前記拡張レイヤ復号化手段は、推定誤差スペクトルまたは誤差スペクトルのいずれかと聴覚マスキングまたは推定聴覚マスキングのいずれかとの振幅値の差を周波数毎に算出し、前記振幅値の差の大きさに基づ

いて復号化の情報量を決定する構成を探る。

【0052】

この構成によれば、拡張レイヤでの符号化において、推定誤差スペクトルが推定聴覚マスキングを超える量に応じて適応ビット配分されたベクトル量子化を行うことにより、量子化効率の向上を図ることができる。

【0053】

本発明の音響復号化装置は、前記拡張レイヤ復号化手段は、前記決定手段で求めた周波数に加えて、予め定められた帯域にある前記誤差スペクトルとを復号化する構成を探る。

【0054】

この構成によれば、あらかじめ定めておいた帯域に含まれるM D C T係数を復号化することにより、符号化の対象として選択されにくいが聴覚的に重要な帯域を強制的に量子化された信号を復号化することができ、符号化側において本来符号化の対象として選択されるべき周波数が選択されない場合でも、聴覚的に重要な帯域に含まれる周波数に位置する誤差スペクトルは必ず量子化されることになり、品質の改善することができる。

【0055】

本発明の音響信号送信装置は、音響信号を電気的信号に変換する音響入力手段と、この音響入力手段から出力された信号をデジタル信号に変換するA／D変換手段と、このA／D変換手段から出力されたデジタル信号を符号化する上記音響符号化装置と、この符号化装置から出力された符号化コードを無線周波数の信号に変調するR F変調手段と、このR F変調手段から出力された信号を電波に変換して送信する送信アンテナとを具備する構成を探る。

【0056】

本発明の音響信号受信装置は、電波を受信する受信アンテナと、この受信アンテナに受信された信号を復調するR F復調手段と、このR F復調手段にて得られた情報を復号する上記音響復号化装置と、この復号化装置から出力された信号をアナログ信号に変換するD／A変換手段と、このD／A変換手段から出力された電気的信号を音響信号に変換する音響出力手段とを具備する構成を探る。

【0057】

本発明の通信端末装置は、上記音響信号送信装置あるいは上記音響信号受信装置の少なくとも一方を具備する構成を探る。本発明の基地局装置は、上記音響信号送信装置あるいは上記音響信号受信装置の少なくとも一方を具備する構成を探る。

【0058】

これらの構成によれば、符号化信号の復号化した信号から拡張レイヤの符号化の対象となる周波数を決定することにより、符号化側から復号化側に伝送する基本レイヤの符号化信号のみで拡張レイヤの符号化の対象となる周波数を決定することができ、符号化側から復号化側にこの周波数の情報を伝送する必要がなくなり、低ビットレートで高品質に符号化を行うことができる。

【0059】

本発明の音響符号化方法は、符号化側において、サンプリングレートが下げられた入力信号を符号化して第1符号化コードを作成し、前記第1符号化コードを復号化した第1復号信号のサンプリングレートを入力時の入力信号のサンプリングレートと同一のレートに上げ、サンプリングレートが上げられた復号信号に基づいて前記誤差信号を符号化する対象の周波数を決定し、入力時の入力信号とサンプリングレートが上げられた復号信号との差分信号のうち前記周波数にある前記差分信号を符号化して第2符号化コードを作成し、復号化側において、前記第1符号化コードを復号化して第2復号信号を得て、前記第2復号信号のサンプリングレートを第3復号信号のサンプリングレートと同一のレートに上げ、このサンプリングレートの上げられた第2復号信号に基づいて第2符号化コードを復号化する対象の周波数を決定し、前記周波数の情報を用いて前記第2符号化コードを復号化して第3復号信号を得て、前記サンプリングレートの上げられた第2復号信号と前記第3復号信号とを加算するようにした。

【0060】

この方法によれば、符号化信号の復号化した信号から拡張レイヤの符号化の対象となる周波数を決定することにより、符号化側から復号化側に伝送する基本レイヤの符号化信号のみで拡張レイヤの符号化の対象となる周波数を決定すること

ができ、符号化側から復号化側にこの周波数の情報を伝送する必要がなくなり、低ビットレートで高品質に符号化を行うことができる。

【0061】

【発明の実施の形態】

本発明者は、入力信号の代わりに基本レイヤの符号化コードを復号化した信号を使って拡張レイヤで符号化する周波数を推定しても、この復号信号は入力信号との歪が小さくなるよう決定されているため、充分に近似され大きな問題は生じないということに着目し本発明をするに至った。

【0062】

本発明の骨子は、入力信号をダウンサンプリングして符号化し、符号化した信号を復号化してアップサンプリングし、このアップサンプリングした復号信号と入力信号との差分信号を符号化する符号化方法において、符号化側と復号化側の両方で算出されるアップサンプリングした復号信号から拡張レイヤで符号化もしくは復号化の対象となる周波数を決定することにより、符号化側から復号化側にこの周波数の情報を伝送ことなく、低ビットレートで高品質に符号化を行うことである。

【0063】

以下、本発明の実施の形態について図面を参照して詳細に説明する。

(実施の形態1)

図1は、本発明の実施の形態1に係る音響符号化装置の構成を示すブロック図である。図1の音響符号化装置100は、ダウンサンプリング器101と、基本レイヤ符号化器102と、局所復号化器103と、アップサンプリング器104と、遅延器105と、減算器106と、周波数決定部107と、拡張レイヤ符号化器108と、多重化器109とから主に構成される。

【0064】

図1において、ダウンサンプリング器101は、サンプリングレートFHの入力データ(音響データ)を受けつけ、この入力データをサンプリングレートFLより低いサンプリングレートFLに変換して基本レイヤ符号化器102に出力する。

【0065】

基本レイヤ符号化器102は、サンプリングレートFLの入力データを所定の基本フレーム単位で符号化し、入力データを符号化した第1符号化コードを局所復号化器103と多重化器109に出力する。例えば、基本レイヤ符号化器102は、入力データをCELP方式で符号化する。

【0066】

局所復号化器103は、第1符号化コードを復号化し、復号化により得られた復号信号をアップサンプリング器104に出力する。アップサンプリング器104は、復号信号のサンプリングレートをFHに上げて減算器106と周波数決定部107に出力する。

【0067】

遅延器105は、入力信号を所定の時間遅延して減算器106に出力する。この遅延の大きさをダウンサンプリング器101と基本レイヤ符号化器102と局所復号化器103とアップサンプリング器104で生じる時間遅れと同値とすることにより、次の減算処理での位相のずれを防ぐ役割を持つ。例えば、この遅延時間は、ダウンサンプリング器101、基本レイヤ符号化器102、局所復号化器103、及びアップサンプリング器104における処理の時間の総和とする。減算器106は、入力信号を復号信号で減算し、減算結果を誤差信号として拡張レイヤ符号化器108に出力する。

【0068】

周波数決定部107は、サンプリングレートをFHに上げた復号信号から誤差信号の符号化する領域と、符号化しない領域を決定して拡張レイヤ符号化器108に通知する。例えば、周波数決定部107は、サンプリングレートをFHに上げた復号信号から聴覚マスキングする周波数を決定して拡張レイヤ符号化器108に出力する。

【0069】

拡張レイヤ符号化器108は、誤差信号を周波数領域の係数に変換して誤差スペクトルを生成し、周波数決定部107から得られる符号化の対象となる周波数情報に基づき誤差スペクトルの符号化を行う。多重化器109は、基本レイヤ符

号化器102で符号化された信号と、拡張レイヤ符号化器108で符号化された信号を多重化する。

【0070】

以下、基本レイヤ符号化器102と拡張レイヤ符号化器108とがそれぞれ符号化する信号について説明する。図2は、音響信号の情報の分布の一例を示す図である。図2において、縦軸は情報量を示し、横軸は周波数を示す。図2では、入力信号に含まれる音声情報と背景音楽・背景雑音情報がどの周波数帯にどれだけ存在しているかを表している。

【0071】

図2に示すように、音声情報は、周波数の低い領域に情報が多く存在し、高域に向かうほど情報量は減少する。一方、背景音楽・背景雑音情報は、音声情報と比べると相対的に低域の情報は少なく、高域に含まれる情報が大きい。

【0072】

そこで、基本レイヤではC E L Pを用いて音声信号を高品質に符号化し、拡張レイヤでは基本レイヤで表しきれない背景の音楽や環境音、基本レイヤでカバーする周波数帯よりも高い周波数成分の信号を効率よく符号化する。

【0073】

図3は、基本レイヤと拡張レイヤで符号化の対象とする領域の一例を示す図である。図3において、縦軸は情報量を示し、横軸は周波数を示す。図3は、基本レイヤ符号化器102と拡張レイヤ符号化器108がそれぞれ符号化する情報の対象となる領域を表している。

【0074】

基本レイヤ符号化器102は、0～F L間の周波数帯の音声情報を効率よく表すように設計されており、この領域での音声情報は品質良く符号化することができる。しかし、基本レイヤ符号化器102では、0～F L間の周波数帯の背景音楽・背景雑音情報の符号化品質が高くない。

【0075】

拡張レイヤ符号化器108は、上記説明にある基本レイヤ符号化器102の能力不足の部分と、F L～F H間の周波数帯の信号をカバーするように設計されて

いる。よって、基本レイヤ符号化器102と拡張レイヤ符号化器108を組み合わせることで広い帯域で高品質な符号化が実現できる。

【0076】

図3に示すように、基本レイヤ符号化器102における符号化により得られた第1符号化コードには、0～FL間の周波数帯の音声情報が含まれているので、少なくとも第1符号化コードのみでも復号信号が得られるというスケーラブル機能が実現できる。

【0077】

また、拡張レイヤで聴覚マスキングを利用して符号化効率を上げることが考えられる。聴覚マスキングとは、ある信号が与えられたときその信号の周波数の近傍に位置する信号が聞こえなくなる（マスクされる）という人間の聴覚特性を利用したものである。

【0078】

図28は、音響（音楽）信号のスペクトルの一例を示す図である。図28において、実線は聴覚マスキングを表し、破線は誤差スペクトルを表している。ここでいう誤差スペクトルとは、入力信号と基本レイヤの復号信号との誤差信号（拡張レイヤの入力信号）のスペクトルを指す。

【0079】

図28の斜線部で表される誤差スペクトルは、聴覚マスキングよりも振幅値が小さいため人間の聴覚では聞こえず、それ以外の領域では誤差スペクトルの振幅値が聴覚マスキングを超えてるので量子化歪が知覚される。

【0080】

そこで、拡張レイヤでは図28の白地部に含まれる誤差スペクトルを符号化してその領域の量子化歪が聴覚マスキングよりも小さくなるようにすればよい。また、斜線部に属する係数は既に聴覚マスキングよりも小さくので量子化する必要がない。

【0081】

本実施の形態の音響符号化装置100では、聴覚マスキング等により残差信号を符号化する周波数を符号化側から復号化側に伝送することをせず、符号化側と

復号側でそれぞれアップサンプリングされた基本レイヤの復号信号を用いて拡張レイヤが符号化する誤差スペクトルの周波数を決定する。

【0082】

基本レイヤの符号化コードを復号化した復号信号は、符号化側と復号化側で同じ信号が得られるので、符号化側は、この復号化信号から聴覚マスキングする周波数を決定して信号を符号化し、復号化側は、この復号化信号から聴覚マスキングされた周波数の情報を得て信号を復号化することにより、誤差スペクトルの周波数の情報を付加情報として符号化して伝送する必要は無くなり、ビットレートの低減を実現することができる。

【0083】

次に、本実施の形態に係る音響符号化装置の各ブロックの詳細な動作について説明する。最初にアップサンプリングされた基本レイヤの復号信号（以後、基本レイヤ復号信号と呼ぶ）から拡張レイヤにて符号化される誤差スペクトルの周波数を決定する周波数決定部107の動作の説明を行う。図4は、本実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブロック図である。

【0084】

図4において、周波数決定部107は、FFT部401と、推定聴覚マスキング算出器402と、決定部403とから主に構成される。

【0085】

FFT部401は、アップサンプリング器104から出力された基本レイヤ復号信号 $x(n)$ を直交変換して振幅スペクトル $P(m)$ を算出して推定聴覚マスキング算出器402と決定部403に出力する。具体的には、FFT部401は、以下の式(1)を用いて振幅スペクトル $P(m)$ を算出する。

【0086】

【数1】

$$P(m) = \sqrt{\operatorname{Re}^2(m) + \operatorname{Im}^2(m)}$$

ここで、 $\operatorname{Re}(m)$ と $\operatorname{Im}(m)$ は基本レイヤ復号信号 $x(n)$ のフーリエ係数の実部と虚部、 m は周波数を表す。

【0087】

次に、推定聴覚マスキング算出器402は、基本レイヤ復号信号の振幅スペクトルP (m) を用いて推定聴覚マスキングM' (m) を算出して決定部403に出力する。一般的には、聴覚マスキングは、入力信号のスペクトルを基に算出されるものであるが、本実施例では入力信号の代わりに基本レイヤ復号信号x (n) を使って聴覚マスキングを推定する。これは、基本レイヤ復号信号x (n) は入力信号との歪が小さくなるよう決定されているため、入力信号の代わりに基本レイヤ復号信号x (n) を用いても充分に近似され大きな問題は生じないという考えに基づいている。

【0088】

次に、決定部403は、基本レイヤ復号信号の振幅スペクトルP (m) と推定聴覚マスキング算出器402で得られる推定聴覚マスキングM' (m) を用いて拡張レイヤ符号化器108で誤差スペクトルを符号化する対象の周波数を決定する。決定部403は、基本レイヤ復号信号の振幅スペクトルP (m) を誤差スペクトルの近似値とみなし、次の式(2)が成り立つ周波数mを拡張レイヤ符号化器108に出力する。

【0089】

【数2】

$$P(m) - M'(m) > 0$$

【0090】

式(2)において、P (m) の項は、誤差スペクトルの大きさを推定しており、M' (m) の項は、聴覚マスキングを推定している。そして、決定部403は、推定誤差スペクトルと推定聴覚マスキングの大きさを比較し、式(2)を満たす場合、すなわち推定聴覚マスキングの大きさを推定誤差スペクトルの大きさが超える場合に、その周波数の誤差スペクトルはノイズとして知覚されるとして拡張レイヤ符号化器108で符号化する対象とする。

【0091】

逆に推定聴覚マスキングの大きさより推定誤差スペクトルの大きさが下回る場

合に、決定部403は、マスキング効果によりその周波数の誤差スペクトルはノイズとして知覚されないとみなし、この周波数の誤差スペクトルは量子化の対象から外す。

【0092】

次に、推定聴覚マスキング算出器402の動作を説明する。図5は、本実施の形態の音響符号化装置の聴覚マスキング算出器の内部構成の一例を示す図である。図5において、推定聴覚マスキング算出器402は、バーカスペクトル算出器501と、スプレッド関数量み込み器502と、トーナリティ算出器503と、聴覚マスキング算出器504とから主に構成される。

【0093】

図5において、バーカスペクトル算出器501は、以下の式(3)を用いてバーカスペクトルB(k)を算出する。

【0094】

【数3】

$$B(k) = \sum_{m=f_L(k)}^{f_H(k)} P^2(m)$$

ここで、P(m)は振幅スペクトルを表し、上述の式(1)より求められる。また、kはバーカスペクトルの番号に対応し、FL(k)、FH(k)はそれぞれ第kバーカスペクトルの最低周波数、最高周波数を表す。バーカスペクトルB(k)はバーカスケール上で等間隔に帯域分割されたときのスペクトル強度を表す。ヘルツスケールをf、バーカスケールをBと表したとき、ヘルツスケールとバーカスケールの関係は以下の式(4)で表される。

【0095】

【数4】

$$B = 13 \tan^{-1}(0.76f) + 3.5 \tan^{-1}\left(\frac{f}{7.5}\right)$$

【0096】

スプレッド関数量み込み器502は、以下に示す式(5)を用いてバーカスペクトルB(k)にスプレッド関数SF(k)を畳み込み、C(k)を算出する。

【0097】

【数5】

$$C(k) = B(k) * SF(k)$$

【0098】

トーナリティ算出器503は、以下の式(6)を用い、各バーチスペクトルのスペクトル平坦度SFM(k)を求める。

【0099】

【数6】

$$SFM(k) = \frac{\mu g(k)}{\mu a(k)}$$

ここで、 $\mu g(k)$ は第kバーチスペクトルに含まれるパワースペクトルの幾何平均、 $\mu a(k)$ は第kバーチスペクトルに含まれるパワースペクトルの算術平均を表す。そして、トーナリティ算出器503は、以下の式(7)を用いてスペクトル平坦度SFM(k)のデシベル値SFMdB(k)からトーナリティ係数 $\alpha(k)$ を算出する。

【0100】

【数7】

$$\alpha(k) = \min\left(\frac{SFMdB(k)}{-60}, 1.0\right)$$

【0101】

聴覚マスキング算出器504は、以下の式(8)を用いてトーナリティ算出器503で算出したトーナリティ係数 $\alpha(k)$ から各バーチスケールのオフセット $O(k)$ を求める。

【0102】

【数8】

$$O(k) = \alpha(k) \cdot (14.5 - k) + (1.0 - \alpha(k)) \cdot 5.5$$

【0103】

そして、聴覚マスキング算出器504は、以下の式(9)を用いてスプレッド関数畳み込み器502で求めたC(k)からオフセットO(k)を減算して聴覚マスキングT(k)を算出する。

【0104】

【数9】

$$T(k) = \max\left(10^{\log_{10}(C(k)) - (O(k)/10)}, T_q(k)\right)$$

ここで、 $T_q(k)$ は絶対閾値を表す。絶対閾値は、人間の聴覚特性として観測される聴覚マスキングの最小値を表す。そして、聴覚マスキング算出器504は、バーカスケールで表される聴覚マスキングT(k)をヘルツスケールに変換して推定聴覚マスキングM'(m)を求め、決定部403に出力する。

【0105】

このようにして求められた量子化の対象となる周波数mを使って、拡張レイヤ符号化器108にてMDCT係数の符号化を行う。図6は、本実施の形態の拡張レイヤ符号化器の内部構成の一例を示すブロック図である。図6の拡張レイヤ符号化器108は、MDCT部601と、MDCT係数量子化器602とから主に構成される。

【0106】

MDCT部601は、減算器106から出力された入力信号に分析窓を乗じた後、MDCT変換(変形離散コサイン変換)してMDCT係数を求める。MDCT変換は、前後の隣接フレームと分析フレームを半分ずつ完全に重ね合わせ、分析フレームの前半部は奇関数、後半部は偶関数という直交基底を用いる。MDCT変換は、波形を合成する際、逆変換後の波形を重ね合わせて加算することにより、フレーム境界歪が発生しないという特徴がある。MDCTを行う際には、sin窓などの窓関数を入力信号に乘ずる。MDCT係数をX(n)とすると、MDCT係数は、式(10)に従い算出される。

【0107】

【数10】

$$X(m) = \sqrt{\frac{1}{N}} \sum_{n=0}^{2N-1} x(n) \cos \left\{ \frac{(2n+1+N) \cdot (2m+1)\pi}{4N} \right\}$$

【0108】

MDCT係数量子化器602は、MDCT部601から出力された入力信号に周波数決定部107から出力された量子化の対象となる周波数に対応する係数を量子化する。そして、MDCT係数量子化器602は、量子化したMDCT係数の符号化コードを多重化器109に出力する。

【0109】

このように、本実施の形態の音響符号化装置によれば、基本レイヤの符号化コードを復号化した信号から拡張レイヤの符号化の対象となる周波数を決定することにより、符号化側から復号化側に伝送する基本レイヤの符号化信号のみで拡張レイヤの符号化の対象となる周波数を決定することができ、符号化側から復号化側にこの周波数の情報を伝送する必要がなくなり、低ビットレートで高品質に符号化を行うことができる。

【0110】

なお、上記実施の形態では、FFTを使った聴覚マスキングの算出法について説明しているが、FFTの代わりMDCTを使って聴覚マスキングを算出することもできる。図7は、本実施の形態の周波数決定部の内部構成の一例を示すブロック図である。但し、図5と同一の構成となるものについては、図5と同一番号を付し、詳しい説明を省略する。

【0111】

MDCT部701は、MDCT係数を使って振幅スペクトルP(m)を近似する。具体的には、MDCT部701は、以下の式(11)を用いてP(m)を近似する。

【0112】

【数11】

$$P(m) = \sqrt{R^2(m)}$$

ここで、R (m) は、アップサンプリング器 104 から与えられる信号をMDC T変換して求めたMDC T係数を表す。

【0113】

推定聴覚マスキング算出器 402 は、MDC T部 701 において近似された P (m) からバーカスペクトル B (k) を算出する。それ以後は上述した方法に従い量子化の対象となる周波数情報を算出する。

【0114】

このように、本実施の形態の音響符号化装置は、MDC Tを使って聴覚マスキングを算出することもできる。

【0115】

次に、復号化側について説明する。図 8 は、本発明の実施の形態 1 に係る音響復号化装置の構成を示すブロック図である。図 8 の音響復号化装置 800 は、分離器 801 と、基本レイヤ復号化器 802 と、アップサンプリング器 803 と、周波数決定部 804 と、拡張レイヤ復号化器 805 と、加算器 806 とから主に構成される。

【0116】

分離器 801 は、音響符号化装置 100 において符号化されたコードを基本レイヤ用の第 1 符号化コードと拡張レイヤ用の第 2 符号化コードに分離し、第 1 符号化コードを基本レイヤ復号化器 802 に出力し、第 2 符号化コードを拡張レイヤ復号化器 805 に出力する。

【0117】

基本レイヤ復号化器 802 は、第 1 符号化コードを復号してサンプリングレート F L の復号信号を得る。そして、基本レイヤ復号化器 802 は、復号信号をアップサンプリング器 803 に出力する。アップサンプリング器 803 は、サンプリングレート F L の復号信号をサンプリングレート F H の復号信号に変換して周波数決定部 804 と加算器 806 に出力する。

【0118】

周波数決定部 804 は、アップサンプリングされた基本レイヤの復号信号を用いて拡張レイヤ復号化器 805 で復号化の対象となる誤差スペクトルの周波数を

決定する。この周波数決定部804は、図1の周波数決定部107と同様の構成をとる。

【0119】

拡張レイヤ復号化器805は、第2符号化コードを復号してサンプリングレートF_Hの復号信号を得る。そして、拡張レイヤ復号化器805は、復号された拡張フレーム単位の復号信号を重ね合わせ、重ね合わせた復号信号を加算器806に出力する。具体的には、拡張レイヤ復号化器805は、復号信号に合成用の窓関数を乗じ、前フレームで復号された時間領域の信号とフレームの半分だけオーバーラップさせて加算して出力信号を生成する。

【0120】

加算器806は、アップサンプリング器803においてアップサンプリングされた基本レイヤの復号信号と、加算器806において復号化された拡張レイヤの復号信号とを加算して出力する。

【0121】

次に、本実施の形態に係る音響復号化装置の各ブロックの詳細な動作について説明する。図9は、本実施の形態の音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図である。図9は、図8の拡張レイヤ復号化器805の内部構成の一例を示す図である。図9の拡張レイヤ復号化器805は、MDCT係数復号化器901と、IMDCT部902と、重ね合わせ加算器903とから主に構成される。

【0122】

MDCT係数復号化器901は、周波数決定部804から出力される復号化の対象となる誤差スペクトルの周波数に基づいて分離器801から出力される第2符号化コードから量子化されたMDCT係数を復号する。具体的には、周波数決定部804から示された信号の周波数に対応する復号MDCT係数を配置し、それ以外の周波数にはゼロを与える。

【0123】

IMDCT部902は、MDCT係数復号化器901から出力されるMDCT係数に逆MDCT変換を施し、時間領域の信号を生成して重ね合わせ加算器90

3に出力する。

【0124】

重ね合わせ加算器903は、復号された拡張フレーム単位の復号信号を重ね合わせ、重ね合わせた復号信号を加算器806に出力する。具体的には、重ね合わせ加算器903は、復号信号に合成用の窓関数を乗じ、前フレームで復号された時間領域の信号とフレームの半分だけオーバーラップさせて加算して出力信号を生成する。

【0125】

このように、本実施の形態の音響復号化装置によれば、基本レイヤの符号化コードを復号化した信号から拡張レイヤの復号化の対象となる周波数を決定することにより、符号化側から復号化側に伝送する基本レイヤの符号化コードのみで拡張レイヤの復号化の対象となる周波数を決定することができ、符号化側から復号化側にこの周波数の情報を伝送する必要がなくなり、低ビットレートで高品質に符号化を行うことができる。

【0126】

(実施の形態2)

本実施の形態では、基本レイヤの符号化においてC E L Pを用いる例について説明する。図10は、本発明の実施の形態2の基本レイヤ符号化器の内部構成の一例を示すブロック図である。図10は、図1の基本レイヤ符号化器102の内部構成を示す図である。図10の基本レイヤ符号化器102は、LPC分析器1001と、聴感重み部1002と、適応符号帳探索器1003と、適応ゲイン量子化器1004と、目標ベクトル生成器1005と、雑音符号帳探索器1006と、雑音ゲイン量子化器1007と、多重化器1008とから主に構成される。

【0127】

LPC分析器1001は、サンプリングレートF Lの入力信号のLPC係数を算出し、このLPC係数をLSP係数などの量子化に適したパラメータに変換して量子化する。そして、LPC分析器1001は、この量子化で得られる符号化コードを多重化器1008に出力する。

【0128】

また、LPC分析器1001は、符号化コードから量子化後のLSP係数を算出してLPC係数に変換し、量子化後のLPC係数を、適応符号帳探索器1003、適応ゲイン量子化器1004、雑音符号帳探索器1006、及び雑音ゲイン量子化器1007に出力する。さらに、LPC分析器1001は、量子化前のLPC係数を聴感重み部1002、適応符号帳探索器1003、適応ゲイン量子化器1004、雑音符号帳探索器1006、及び雑音ゲイン量子化器1007に出力する。

【0129】

聴感重み部1002は、LPC分析器1001で求められたLPC係数に基づいてダウンサンプリング器101から出力された入力信号に重み付けを行う。これは、量子化歪のスペクトルを入力信号のスペクトル包絡にマスクされるようスペクトル整形を行うことを目的としている。

【0130】

適応符号帳探索器1003では、聴覚重み付けされた入力信号を目標信号として適応符号帳の探索が行われる。過去の音源系列をピッチ周期で繰り返した信号を適応ベクトルと呼び、あらかじめ定められた範囲のピッチ周期で生成された適応ベクトルによって適応符号帳は構成される。

【0131】

聴覚重み付けされた入力信号を $t(n)$ 、ピッチ周期 i の適応ベクトルに量子化前のLPC係数と量子化後のLPC係数で構成される重み付き合成フィルタのインパルス応答を畳み込んだ信号を $p_i(n)$ としたとき、適応符号帳探索器1003は、式(12)の評価関数Dを最小とする適応ベクトルのピッチ周期 i をパラメータとして多重化器1008に出力する。

【0132】

【数12】

$$D = \sum_{n=0}^{N-1} t^2(n) - \frac{\left(\sum_{n=0}^{N-1} t(n)p_i(n) \right)^2}{\sum_{n=0}^{N-1} p_i^2(n)}$$

ここで、Nはベクトル長を表す。式(12)の第1項はピッチ周期 i に独立な

で、実際には、適応符号帳探索器1003は第2項のみを計算する。

【0133】

適応ゲイン量子化器1004は、適応ベクトルに乘じられる適応ゲインの量子化を行う。適応ゲイン β は、以下の式(13)で表され、適応ゲイン量子化器1004は、この適応ゲイン β をスカラー量子化し、量子化時に得られる符号を多重化器1008に出力する。

【0134】

【数13】

$$\beta = \frac{\sum_{n=0}^{N-1} t(n) p_i(n)}{\sum_{n=0}^{N-1} p_i^2(n)}$$

【0135】

目標ベクトル生成器1005は、入力信号から適応ベクトルの影響を減算して、雑音符号帳探索器1006と雑音ゲイン量子化器1007で用いる目標ベクトルを生成して出力する。目標ベクトル生成器1005は、 $p_i(n)$ を式12で表される評価関数Dを最小とするときの適応ベクトルに重み付き合成フィルタのインパルス応答を畳み込んだ信号、 βq を式13で表される適応ベクトル β をスカラー量子化したときの量子化値としたとき、目標ベクトル $t_2(n)$ は、以下に示す式(14)のように表される。

【0136】

【数14】

$$t2(n) = t(n) - \beta q \cdot p_i(n)$$

【0137】

雑音符号帳探索器1006は、前記目標ベクトル $t_2(n)$ と量子化前のLPC係数と量子化後のLPC係数を用いて雑音符号帳の探索を行う。例えば、雑音符号帳探索器1006には、ランダム雑音や大規模な音声信号を使って学習した信号を用いることができる。また、雑音符号帳探索器1006が備える雑音符号

帳は、代数(Algebraic)符号帳のように、振幅1のパルスをあらかじめ定められた非常に少ない数だけ有するベクトルで表されることができる。この代数符号帳は、パルスの位置とパルスの符号(極性)の最適な組み合わせを少ない計算量で決定することができるという特徴がある。

【0138】

雑音符号帳探索器1006は、目標ベクトルを $t_2(n)$ 、コード j に対応する雑音ベクトルに重み付き合成フィルタのインパルス応答を畳み込んだ信号を $c_j(n)$ としたとき、以下に示す式(15)の評価関数Dを最小とする雑音ベクトルのインデックス j を多重化器1008に出力する。

【0139】

【数15】

$$D = \sum_{n=0}^{N-1} t_2^2(n) - \frac{\left(\sum_{n=0}^{N-1} t_2(n)c_j(n) \right)^2}{\sum_{n=0}^{N-1} c_j^2(n)}$$

【0140】

雑音ゲイン量子化器1007は、雑音ベクトルに乘じる雑音ゲインを量子化する。雑音ゲイン量子化器1007は、以下に示す式(16)を用いて雑音ゲイン γ を算出し、この雑音ゲイン γ をスカラー量子化して多重化器1008に出力する。

【0141】

【数16】

$$\gamma = \frac{\sum_{n=0}^{N-1} t_2(n)c_j(n)}{\sum_{n=0}^{N-1} c_j^2(n)}$$

【0142】

多重化器1008は、送られてきたLPC係数、適応ベクトル、適応ゲイン、雑音ベクトル、雑音ゲインの符号化コードを多重化して局所復号化器103及び多重化器109に出力する。

【0143】

次に、復号化側について説明する。図11は、本実施の形態の基本レイヤ復号化器の内部構成の一例を示すブロック図である。図11は、図6の基本レイヤ復号化器802の内部構成を示す図である。図11の基本レイヤ復号化器802は、分離器1101と、音源生成器1102と、合成フィルタ1103とから主に構成される。

【0144】

分離器1101は、分離器801から出力された第1符号化コードをLPC係数、適応ベクトル、適応ゲイン、雑音ベクトル、雑音ゲインの符号化コードに分離して、適応ベクトル、適応ゲイン、雑音ベクトル、雑音ゲインの符号化コードを音源生成器1102に出力する。同様に、分離器1101は、LPC係数の符号化コードを合成フィルタ1103に出力する。

【0145】

音源生成器1102は、適応ベクトル、適応ベクトルゲイン、雑音ベクトル、雑音ベクトルゲインの符号化コードを復号し、以下に示す式(17)を用いて音源ベクトル $ex(n)$ を生成する。

【0146】

【数17】

$$ex(n) = \beta_q \cdot q(n) - \gamma_q \cdot c(n)$$

ここで、 $q(n)$ は適応ベクトル、 β_q は適応ベクトルゲイン、 $c(n)$ は雑音ベクトル、 γ_q は雑音ベクトルゲインを表す。

【0147】

合成フィルタ1103では、LPC係数の符号化コードからLPC係数を復号し、以下に示す式(18)を用いて復号されたLPC係数から合成信号 $syn(n)$ を生成する。

【0148】

【数18】

$$syn(n) = ex(n) + \sum_{i=1}^{NP} \alpha_q(i) \cdot syn(n-i)$$

ここで、 α q は復号された L P C 係数、 N P は L P C 係数の次数を表す。そして、合成フィルタ 1103 は、復号された復号信号 s y n (n) をアップサンプリング器 803 に出力する。

【0149】

このように、本実施の形態の音響符号化装置及び音響復号化装置によれば、送信側において、基本レイヤに C E L P を適用して入力信号を符号化し、受信側において、この符号化した入力信号に C E L P を適用して復号することにより、低ビットレートで高品質な基本レイヤを実現することができる。

【0150】

なお、本実施の形態の音声符号化装置は、量子化歪の知覚を抑制するために、合成フィルタ 1103 の後にポストフィルタを従属接続する構成を探ることもできる。図 12 は、本実施の形態の基本レイヤ復号化器の内部構成の一例を示すブロック図である。但し、図 11 と同一の構成となるものについては、図 11 と同一番号を付し、詳しい説明を省略する。

【0151】

ポストフィルタ 1201 は、量子化歪の知覚の抑制の実現のために様々な構成を適用しうるが、代表的な方法として、分離器 1101 で復号されて得られる L P C 係数から構成されるホルマント強調フィルタを用いる方法がある。ホルマント強調フィルタ $H_f(z)$ は以下に示す式 (19) で表される。

【0152】

【数19】

$$H_f(z) = \frac{A(z/\gamma_n)}{A(z/\gamma_d)} (1 - \mu z^{-1})$$

ここで、 $A(z)$ は復号 L P C 係数から構成される合成フィルタ、 γ_n 、 γ_d 、 μ はフィルタの特性を決定する定数を表す。

【0153】

(実施の形態 3)

図 13 は、本発明の実施の形態 3 に係る音響符号化装置の周波数決定部の内部構成の一例を示すブロック図である。但し、図 4 と同一の構成となるものについ

ては、図4と同一番号を付し、詳しい説明を省略する。図13の周波数決定部107は、推定誤差スペクトル算出器1301と、決定部1302とを具備し、基本レイヤ復号信号の振幅スペクトルP(m)から推定誤差スペクトルE'(m)を推定し、推定誤差スペクトルE'(m)と推定聴覚マスキングM'(m)とを用いて拡張レイヤ符号化器108で符号化される誤差スペクトルの周波数を決定する点が図4と異なる。

【0154】

FFT部401は、アップサンプリング器104から出力された基本レイヤ復号信号x(n)を直交変換して振幅スペクトルP(m)を算出して推定聴覚マスキング算出器402と推定誤差スペクトル算出器1301に出力する。

【0155】

推定誤差スペクトル算出器1301は、FFT部401で算出される基本レイヤ復号信号の振幅スペクトルP(m)から推定誤差スペクトルE'(m)を算出して決定部1302に出力する。推定誤差スペクトルE'(m)は、基本レイヤ復号信号の振幅スペクトルP(m)を平坦に近づける処理を施し算出される。具体的には、推定誤差スペクトル算出器1301は、以下の式(20)を用いて推定誤差スペクトルE'(m)を算出する。

【0156】

【数20】

$$E'(m) = a \cdot P(m)^\gamma$$

ここでaと γ は0以上1未満の定数を表す。

【0157】

決定部1302は、推定誤差スペクトル算出器1301において推定された推定誤差スペクトルE'(m)と推定聴覚マスキング算出器402で得られる推定聴覚マスキングM'(m)を用いて拡張レイヤ符号化器108で誤差スペクトルを符号化する対象の周波数を決定する。

【0158】

次に、本実施の形態の推定誤差スペクトル算出器1301が算出する推定誤差スペクトルについて説明する。図14は、本実施の形態の推定誤差スペクトル算

出器が算出する残差スペクトルの一例を示す図である。

【0159】

誤差スペクトル $E(m)$ は、図14に示すように基本レイヤ復号信号の振幅スペクトル $P(m)$ に比ベスペクトルの形状が平坦になり、かつ全帯域のパワーが小さくなっている。よって、振幅スペクトル $P(m)$ を γ ($0 < \gamma < 1$) 乗することによりスペクトル形状を平坦化し、 a ($0 < a < 1$) 倍することにより全体域のパワーを減少させることにより、誤差スペクトルの推定精度を向上させることができる。

【0160】

同様に、復号化側も音響復号化装置800の周波数決定部804の内部構成を符号化側の図13の周波数決定部107と同じ構成とする。

【0161】

このように、本実施の形態の音響符号化装置によれば、基本レイヤの復号信号のスペクトルから推定した残差スペクトルを平滑化することにより、推定誤差スペクトルを残差スペクトルに近似することができ、拡張レイヤにて誤差スペクトルを効率よく符号化することができる。

【0162】

なお、本実施の形態では、FFTを用いた場合について説明したが、前述した実施の形態1と同様に、FFTの代わりにMDCTを用いる構成も可能である。

【0163】

(実施の形態4)

図15は、本発明の実施の形態4に係る音響符号化装置の周波数決定部の内部構成の一例を示すブロック図である。但し、図4と同一の構成となるものについては、図4と同一番号を付し、詳しい説明を省略する。図15の周波数決定部107は、推定聴覚マスキング修正部1501と、決定部1502とを具備し、周波数決定部107において、基本レイヤ復号信号の振幅スペクトル $P(m)$ から推定聴覚マスキング算出器402にて推定聴覚マスキング $M'(m)$ を算出した後に、この推定聴覚マスキング $M'(m)$ に基本レイヤ符号化器102の符号化コードの情報を基に修正を加える点が図4と異なる。

【0164】

FFT部401は、アップサンプリング器104から出力された基本レイヤ復号信号 $x(n)$ を直交変換して振幅スペクトル $P(m)$ を算出して推定聴覚マスキング算出器402と決定部1502に出力する。推定聴覚マスキング算出器402は、基本レイヤ復号信号の振幅スペクトル $P(m)$ を用いて推定聴覚マスキング $M'(m)$ を算出して推定聴覚マスキング修正部1501に出力する。

【0165】

推定聴覚マスキング修正部1501は、基本レイヤ符号化器102から入力される基本レイヤの符号化コードの情報を用いて推定聴覚マスキング算出器402で求められる推定聴覚マスキング $M'(m)$ に修正を加える。

【0166】

ここでは、基本レイヤの符号化コードの情報として、復号LPC係数から算出される1次のPARCOR係数が与えられるものとする。一般にLPC係数やPARCOR係数は入力信号のスペクトル包絡を表現する。PARCOR係数の次数を下げていくと、PARCOR係数の性質から、スペクトル包絡の形状が簡略化されてゆき、PARCOR係数の次数が1次のときにスペクトルの傾きの程度を表すようになる。

【0167】

一方で、入力信号として与えられる楽音や音声のスペクトル特性には、高域に対して低域にパワーが偏っている場合（例えば母音）やその逆の場合（例えば子音）が存在する。基本レイヤ復号信号はこういった入力信号のスペクトル特性に影響を受けやすく、必要以上にスペクトルのパワーの偏りを強調してしまう傾向にある。

【0168】

そこで、本実施の形態の音響符号化装置は、推定聴覚マスキング修正部1501において、前述した1次のPARCOR係数を利用して過度に強調されたスペクトルの偏りを補正することにより、推定マスキング $M'(m)$ の精度を向上させることができる。

【0169】

推定聴覚マスキング修正部1501は、以下に示す式(21)を用いて基本レイヤ符号化器102から出力された1次のPARCOR係数k(1)から修正フィルタ $H_k(z)$ を算出する。

【0170】

【数21】

$$H_k(z) = 1 - \beta \cdot k(1) \cdot z^{-1}$$

ここで β は1未満の正の定数を表す。次に、推定聴覚マスキング修正部1501は、以下に示す式(22)を用いて $H_k(z)$ の振幅特性 $K(m)$ を算出する。

【0171】

【数22】

$$K(m) = \left| 1 - \beta \cdot k(1) \cdot e^{-j \frac{2\pi m}{M}} \right|$$

【0172】

そして、推定聴覚マスキング修正部1501は、以下の式(23)を用いて修正フィルタの振幅特性 $K(m)$ から修正後の推定聴覚マスキング $M''(m)$ を算出する。

【0173】

【数23】

$$M''(m) = K(m) \cdot M'(m)$$

【0174】

そして、推定聴覚マスキング修正部1501は、推定聴覚マスキング $M'(m)$ の代わりに修正後の聴覚マスキング $M''(m)$ を決定部1502に出力する。

【0175】

決定部1502は、基本レイヤ復号信号の振幅スペクトル $P(m)$ と推定聴覚マスキング修正部1501から出力される修正後の聴覚マスキング $M''(m)$ を用いて拡張レイヤ符号化器108で誤差スペクトルを符号化する対象の周波数

を決定する。

【0176】

このように本実施の形態の音響符号化装置によれば、マスキング効果の特性を利用して、入力信号のスペクトルから聴覚マスキングを算出し、拡張レイヤの符号化において、量子化歪をこのマスキング値以下になるように量子化を行うことにより、品質の劣化を伴わずに量子化の対象となるM D C T係数の数を減らすことができ、低ビットレートで高品質に符号化を行うことができる。

【0177】

このように、本実施の形態の音響符号化装置によれば、基本レイヤ復号信号の振幅スペクトルから推定した推定聴覚マスキングを、基本レイヤ符号化器の符号化コードの情報を基に修正を加えることにより、推定聴覚マスキングの精度を向上させることができ、結果拡張レイヤにて誤差スペクトルを効率よく符号化することができる。

【0178】

同様に、復号化側も音響復号化装置800の周波数決定部804の内部構成を符号化側の図15の周波数決定部107と同じ構成とする。

【0179】

なお、本実施の形態の周波数決定部107は、本実施の形態と実施の形態3とを組み合わせた構成を探ることもできる。図16は、本実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブロック図である。但し、図4と同一の構成となるものについては、図4と同一番号を付し、詳しい説明を省略する。

【0180】

F F T部401は、アップサンプリング器104から出力された基本レイヤ復号信号 $x(n)$ を直交変換して振幅スペクトル $P(m)$ を算出して推定聴覚マスキング算出器402と推定誤差スペクトル算出器1301に出力する。

【0181】

推定聴覚マスキング算出器402は、基本レイヤ復号信号の振幅スペクトル $P(m)$ を用いて推定聴覚マスキング $M'(m)$ を算出して推定聴覚マスキング修

正部1501に出力する。

【0182】

推定聴覚マスキング修正部1501は、基本レイヤ符号化器102から入力される基本レイヤの符号化コードの情報が推定聴覚マスキング修正部1501を用いて推定聴覚マスキング算出器402で求められる推定聴覚マスキングM' (m)に修正を加える。

【0183】

推定誤差スペクトル算出器1301は、FFT部401で算出される基本レイヤ復号信号の振幅スペクトルP (m)から推定誤差スペクトルE' (m)を算出して決定部1601に出力する。

【0184】

決定部1601は、推定誤差スペクトル算出器1301において推定された推定誤差スペクトルE' (m)と推定聴覚マスキング修正部1501から出力される修正後の聴覚マスキングM' (m)を用いて拡張レイヤ符号化器108で誤差スペクトルを符号化する対象の周波数を決定する。

【0185】

また、本実施の形態では、FFTを用いた場合について説明したが、前述した実施の形態1と同様に、FFTの代わりにMDCTを用いる構成も可能である。

【0186】

(実施の形態5)

図17は、本発明の実施の形態5に係る音響符号化装置の拡張レイヤ符号化器の内部構成の一例を示すブロック図である。但し、図6と同一の構成となるものについては、図6と同一番号を付し、詳しい説明を省略する。図17の拡張レイヤ符号化器は、順序づけ部1701と、MDCT係数量子化器1702を具備し、周波数決定部107から与えられる周波数を推定歪値D (m)の大きさに従い周波数別で符号化後の情報量に重み付けを行う点が図6の拡張レイヤ符号化器と異なる。

【0187】

図17において、MDCT部601は、減算器106から出力された入力信号

に分析窓を乗じた後、M D C T 変換(変形離散コサイン変換)してM D C T 係数を求め、M D C T 係数量子化器1702に出力する。

【0188】

順序づけ部1701は、周波数決定部107で求められた周波数情報を受けつけ、各周波数の推定誤差スペクトル E' (m) が推定聴覚マスキング M' (m) を超える量(以後、推定歪値と呼ぶ) D (m) を算出する。この推定歪値 D (m) は、以下に示す式(24)で定義される。

【0189】

【数24】

$$D(m) = E'(m) - M'(m)$$

【0190】

ここで、順序づけ部1701は、以下に示す式(25)を満たす推定歪値 D (m)のみ算出する。

【0191】

【数25】

$$E'(m) - M'(m) > 0$$

【0192】

そして、順序づけ部1701は、推定歪値 D (m) の大きさが大きいものから順序付けを行い、その周波数情報をM D C T 係数量子化器1702に出力する。M D C T 係数量子化器1702では、推定歪値 D (m) により順序付けされた周波数情報を基に、推定歪値 D (m) の大きいものからその周波数に位置する誤差スペクトル E (m) にビットを多く配分して量子化を行う。

【0193】

ここでは例として、周波数決定手段から送られてくる周波数と推定歪値が図18である場合について説明する。図18は、本実施の形態の順序づけ部の推定歪値の順位づけの一例を示す図である。

【0194】

順序づけ部1701は、図18の情報から、推定歪値 D (m) が大きい順に周波数の並べ替えを行う。この例では、順序づけ部1701の処理の結果、周波数

$m = 7, 8, 4, 9, 1, 11, 3, 12$ の順序が得られる。順序づけ部 1701 は、この順序付けの情報をMDCT係数量子化器 1702 に出力する。

【0195】

MDCT係数量子化器 1702 では、MDCT部 601 から与えられる誤差スペクトル $E(m)$ の内、順序づけ部 1701 から与えられる順序付けの情報を基にして、 $E(7), E(8), E(4), E(9), E(1), E(11), E(3), E(12)$ を量子化する。

【0196】

このとき、順序づけの先頭に位置する誤差スペクトルの量子化に用いられるビット数は多く配分され、末尾にいくに従いビット数は少なく配分される。すなわち、推定歪値 $D(m)$ が大きい周波数ほど誤差スペクトルの量子化に用いられるビット数は多く配分され、推定歪値 $D(m)$ が小さい周波数ほど誤差スペクトルの量子化に用いられるビット数は少なく配分される。

【0197】

例えば、 $E(7)$ を 8 ビット、 $E(8), E(4)$ を 7 ビット、 $E(9), E(1)$ を 6 ビット、 $E(11), E(3), E(12)$ を 5 ビットというようなビット配分を実施する。このような推定歪値 $D(m)$ に応じた適応ビット配分を行うことにより、量子化の効率が向上することになる。

【0198】

ベクトル量子化を適用する場合には、拡張レイヤ符号化器 108 は、先頭に位置する誤差スペクトルから順にベクトルを構成し、それぞれのベクトルに対してベクトル量子化を行う。このとき、先頭に位置する誤差スペクトルのビット配分が多くなり、末尾に位置する誤差スペクトルのビット配分が小さくなるようなベクトルの構成と量子化ビット配分が成される。図 18 の例では、 $V1 = (E(7), E(8)), V2 = (E(4), E(9)), V3 = (E(1), E(11), E(3), E(12))$ のように、2 次元、2 次元、4 次元の 3 ベクトルを構成し、 $V1$ を 10 ビット、 $V2$ を 8 ビット、 $V3$ を 8 ビットというようなビット配分を行う。

【0199】

このように、本実施の形態の音響符号化装置によれば、拡張レイヤでの符号化において、推定誤差スペクトルが推定聴覚マスキングを超える量が大きい周波数に多くの情報量を配分して符号化することにより、量子化効率の向上を図ることができる。

【0200】

次に復号化側について説明する。図19は、本発明の実施の形態5に係るの音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図である。但し、図9と同一の構成となるものについては、図9と同一番号を付し、詳しい説明を省略する。図19の拡張レイヤ復号化器805は、順序づけ部1901と、MDCT係数復号化器1902とを具備し、周波数決定部804から与えられる周波数を推定歪値D(m)の大きさに従い順序付けを行う点が図9と異なる。

【0201】

順序づけ部1901は、上述の式(24)を用いて推定歪値D(m)を算出する。順序づけ部1901は、上述の順序づけ部1701と同一の構成を探る。この構成により適応ピット配分を行い量子化効率の向上を図ることができる上述の音響符号化法の符号化コードを復号することができる。

【0202】

MDCT係数復号化器1902は、推定歪値D(m)の大きさに従い順序付けされた周波数の情報を用いて分離器801から出力された第2符号化コードを復号化する。具体的には、MDCT係数復号化器1902は、周波数決定部804から与えられる周波数に対応する復号MDCT係数を配置し、それ以外の周波数にはゼロを与える。次にIMDCT部902は、MDCT係数復号化器1902から得られるMDCT係数に逆MDCT変換を施し、時間領域の信号を生成する。

【0203】

重ね合わせ加算器903は、前記信号に合成用の窓関数を乗じ、前フレームで復号された時間領域の信号とフレームの半分だけオーバーラップさせて加算して出力信号を生成する。重ね合わせ加算器903は、この出力信号を加算器806に出力する。

【0204】

このように、本実施の形態の音響復号化装置によれば、拡張レイヤでの符号化において、推定誤差スペクトルが推定聴覚マスキングを超える量に応じて適応ビット配分されたベクトル量子化を行うことにより、量子化効率の向上を図ることができる。

【0205】

(実施の形態6)

図20は、本発明の実施の形態6に係る音響符号化装置の拡張レイヤ符号化器の内部構成の一例を示すブロック図である。但し、図6と同一の構成となるものについては、図6と同一番号を付し、詳しい説明を省略する。図20の拡張レイヤ符号化器は、固定帯域指定部2001と、MDCT係数量子化器2002とを具備し、あらかじめ定めておいた帯域に含まれるMDCT係数を周波数決定部107から得られる周波数と共に量子化する点が図6の拡張レイヤ符号化器と異なる。

【0206】

図20において、固定帯域指定部2001には、あらかじめ聴感上重要な帯域が設定されている。ここでは、設定されている帯域に含まれる周波数を $m = 15, 16$ とする。

【0207】

MDCT係数量子化器2002は、MDCT部601から出力された入力信号に周波数決定部107から出力された聴覚マスキングを用いて入力信号を量子化する係数と量子化しない係数に分類し、量子化する係数と、さらに固定帯域指定部2001が設定する帯域にある係数を符号化する。

【0208】

その周波数が図18で示されたものであるとすると、MDCT係数量子化器2002では、誤差スペクトルE(1)、E(3)、E(4)、E(7)、E(8)、E(9)、E(11)、E(12)および、固定帯域指定部2001で指定される周波数の誤差スペクトルE(15)、E(16)が量子化される。

【0209】

このように、本実施の形態の音響符号化装置によれば、符号化の対象として選択されにくいが聴覚的に重要な帯域を強制的に量子化することにより、本来符号化の対象として選択されるべき周波数が選択されない場合でも、聴覚的に重要な帯域に含まれる周波数に位置する誤差スペクトルは必ず量子化されることになり、品質を改善することができる。

【0210】

次に、復号化側について説明する。図21は、本発明の実施の形態6に係る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図である。但し、図9と同一の構成となるものについては、図9と同一番号を付し、詳しい説明を省略する。図21の拡張レイヤ復号化器は、固定帯域指定部2101と、MDCT係数復号化器2102とを具備し、あらかじめ定めておいた帯域に含まれるMDCT係数を周波数決定部804から得られる周波数と共に復号化する点が図9の拡張レイヤ復号化器と異なる。

【0211】

図21において、固定帯域指定部2101には、あらかじめ聴感上重要な帯域が設定されている。

【0212】

MDCT係数復号化器2102は、周波数決定部804から出力される復号化の対象となる誤差スペクトルの周波数に基づいて分離器801から出力される第2符号化コードから量子化されたMDCT係数を復号する。具体的には、周波数決定部804と固定帯域指定部2101から示された信号の周波数に対応する復号MDCT係数を配置し、それ以外の周波数にはゼロを与える。

【0213】

IMDCT部902は、MDCT係数復号化器2102から出力されるMDCT係数に逆MDCT変換を施し、時間領域の信号を生成して重ね合わせ加算器903に出力する。

【0214】

このように、本実施の形態の音響復号化装置によれば、あらかじめ定めておいた帯域に含まれるMDCT係数を復号化することにより、符号化の対象として選

択されにくいが聴覚的に重要な帯域を強制的に量子化された信号を復号化することができ、符号化側において本来符号化の対象として選択されるべき周波数が選択されない場合でも、聴覚的に重要な帯域に含まれる周波数に位置する誤差スペクトルは必ず量子化されることになり、品質の改善することができる。

【0215】

なお、本実施の形態の拡張レイヤ符号化器及び拡張レイヤ復号化器は、本実施の形態と実施の形態5とを組み合わせた構成を探ることもできる。図22は、本実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブロック図である。但し、図6と同一の構成となるものについては、図6と同一番号を付し、詳しい説明を省略する。

【0216】

図22において、MDCT部601は、減算器106から出力された入力信号に分析窓を乗じた後、MDCT変換(変形離散コサイン変換)してMDCT係数を求め、MDCT係数量子化器2201に出力する。

【0217】

順序づけ部1701は、周波数決定部107で求められた周波数情報を受けつけ、各周波数の推定誤差スペクトルE' (m) が推定聴覚マスキングM' (m) を超える量(以後、推定歪値と呼ぶ) D (m) を算出する。

固定帯域指定部2001には、あらかじめ聴感上重要な帯域が設定されている。

【0218】

MDCT係数量子化器2201では、推定歪値D (m) により順序付けされた周波数情報を基に、推定歪値D (m) の大きいものからその周波数に位置する誤差スペクトルE (m) にビットを多く配分して量子化を行う。また、MDCT係数量子化器2201は、固定帯域指定部2001が設定する帯域にある係数を符号化する。

【0219】

次に復号化側について説明する。図23は、本発明の実施の形態6に係る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図である。但し、図9と同一の構成となるものについては、図9と同一番号を付し、詳しい説

明を省略する。

【0220】

図23において、順序づけ部1901は、周波数決定部804で求められた周波数情報を受けつけ、各周波数の推定誤差スペクトルE' (m) が推定聴覚マスキングM' (m) を超える量（以後、推定歪値と呼ぶ）D (m) を算出する。

【0221】

そして、順序づけ部1901は、推定歪値D (m) の大きさが大きいものから順序付けを行い、その周波数情報をMDCT係数復号化器2301に出力する。固定帯域指定部2101には、あらかじめ聴感上重要な帯域が設定されている。

【0222】

MDCT係数復号化器2301は、順序づけ部1901から出力される復号化の対象となる誤差スペクトルの周波数に基づいて分離器801から出力される第2符号化コードから量子化されたMDCT係数を復号する。具体的には、周波数決定部804と固定帯域指定部2101から示された信号の周波数に対応する復号MDCT係数を配置し、それ以外の周波数にはゼロを与える。

【0223】

IMDCT部902は、MDCT係数復号化器2301から出力されるMDCT係数に逆MDCT変換を施し、時間領域の信号を生成して重ね合わせ加算器903に出力する。

【0224】

（実施の形態7）

次に、本発明の実施の形態7について、図面を参照して説明する。図24は、本発明の実施の形態7に係る通信装置の構成を示すブロック図である。図23における信号処理装置2403は前述した実施の形態1から実施の形態6に示した音響符号化装置の中の1つによって構成されている点に本実施の形態の特徴がある。

【0225】

図24に示すように、本発明の実施の形態7に係る通信装置2400は、入力装置2401、A/D変換装置2402及びネットワーク2404に接続されて

いる信号処理装置2403を具備している。

【0226】

A/D変換装置2402は、入力装置2401の出力端子に接続されている。信号処理装置2403の入力端子は、A/D変換装置2402の出力端子に接続されている。信号処理装置2403の出力端子はネットワーク2404に接続されている。

【0227】

入力装置2401は、人間の耳に聞こえる音波を電気的信号であるアナログ信号に変換してA/D変換装置2402に与える。A/D変換装置2402はアナログ信号をデジタル信号に変換して信号処理装置2403に与える。信号処理装置2403は入力されてくるデジタル信号を符号化してコードを生成し、ネットワーク2404に出力する。

【0228】

このように、本発明の実施の形態の通信装置によれば、通信において前述した実施の形態1～6に示したような効果を享受でき、少ないビット数で効率よく音響信号を符号化する音響符号化装置を提供することができる。

【0229】

(実施の形態8)

次に、本発明の実施の形態8について、図面を参照して説明する。図25は、本発明の実施の形態8に係る通信装置の構成を示すブロック図である。図25における信号処理装置2503は前述した実施の形態1から実施の形態6に示した音響復号化装置の中の1つによって構成されている点に本実施の形態の特徴がある。

【0230】

図25に示すように、本発明の実施の形態8に係る通信装置2500は、ネットワーク2501に接続されている受信装置2502、信号処理装置2503、及びD/A変換装置2504及び出力装置2505を具備している。

【0231】

受信装置2502の入力端子は、ネットワーク2501に接続されている。信

号処理装置2503の入力端子は、受信装置2502の出力端子に接続されている。D/A変換装置2504の入力端子は、信号処理装置2503の出力端子に接続されている。出力装置2505の入力端子は、D/A変換装置2504の出力端子に接続されている。

【0232】

受信装置2502は、ネットワーク2501からのデジタルの符号化音響信号を受けてデジタルの受信音響信号を生成して信号処理装置2503に与える。信号処理装置2503は、受信装置2502からの受信音響信号を受けてこの受信音響信号に復号化処理を行ってデジタルの復号化音響信号を生成してD/A変換装置2504に与える。D/A変換装置2504は、信号処理装置2503からのデジタルの復号化音声信号を変換してアナログの復号化音声信号を生成して出力装置2505に与える。出力装置2505は、電気的信号であるアナログの復号化音響信号を空気の振動に変換して音波として人間の耳に聴こえるように出力する。

【0233】

このように、本実施の形態の通信装置によれば、通信において前述した実施の形態1~6に示したような効果を享受でき、少ないビット数で効率よく符号化された音響信号を復号することができるので、良好な音響信号を出力することができる。

【0234】

(実施の形態9)

次に、本発明の実施の形態9について、図面を参照して説明する。図26は、本発明の実施の形態9に係る通信装置の構成を示すブロック図である。本発明の実施の形態9において、図26における信号処理装置2603は、前述した実施の形態1から実施の形態6に示した音響符号化器の中の1つによって構成されている点に本実施の形態の特徴がある。

【0235】

図26に示すように、本発明の実施の形態9に係る通信装置2600は、入力装置2601、A/D変換装置2602、信号処理装置2603、RF変調装置

2604及びアンテナ2605を具備している。

【0236】

入力装置2601は人間の耳に聞こえる音波を電気的信号であるアナログ信号に変換してA/D変換装置2602に与える。A/D変換装置2602はアナログ信号をデジタル信号に変換して信号処理装置2603に与える。信号処理装置2603は入力されてくるデジタル信号を符号化して符号化音響信号を生成し、RF変調装置2604に与える。RF変調装置2604は、符号化音響信号を変調して変調符号化音響信号を生成し、アンテナ2605に与える。アンテナ2605は、変調符号化音響信号を電波として送信する。

【0237】

このように、本実施の形態の通信装置によれば、無線通信において前述した実施の形態1～6に示したような効果を享受でき、少ないビット数で効率よく音響信号を符号化することができる。

【0238】

なお、本発明は、オーディオ信号を用いる送信装置、送信符号化装置又は音響信号符号化装置に適用することができる。また、本発明は、移動局装置又は基地局装置にも適用することができる。

【0239】

(実施の形態10)

次に、本発明の実施の形態10について、図面を参照して説明する。図27は、本発明の実施の形態10に係る通信装置の構成を示すブロック図である。本発明の実施の形態10において、図27における信号処理装置2703は、前述した実施の形態1から実施の形態6に示した音響復号化器の中の1つによって構成されている点に本実施の形態の特徴がある。

【0240】

図27に示すように、本発明の実施の形態10に係る通信装置2700は、アンテナ2701、RF復調装置2702、信号処理装置2703、D/A変換装置2704及び出力装置2705を具備している。

【0241】

アンテナ2701は、電波としてのデジタルの符号化音響信号を受けて電気信号のデジタルの受信符号化音響信号を生成してRF復調装置2702に与える。RF復調装置2702は、アンテナ2701からの受信符号化音響信号を復調して復調符号化音響信号を生成して信号処理装置2703に与える。

【0242】

信号処理装置2703は、RF復調装置2702からのデジタルの復調符号化音響信号を受けて復号化処理を行ってデジタルの復号化音響信号を生成してD/A変換装置2704に与える。D/A変換装置2704は、信号処理装置2703からのデジタルの復号化音声信号を変換してアナログの復号化音声信号を生成して出力装置2705に与える。出力装置2705は、電気的信号であるアナログの復号化音声信号を空気の振動に変換して音波として人間の耳に聴こえるように出力する。

【0243】

このように、本実施の形態の通信装置によれば、無線通信において前述した実施の形態1～6に示したような効果を享受でき、少ないビット数で効率よく符号化された音響信号を復号することができるので、良好な音響信号を出力することができる。

【0244】

なお、本発明は、オーディオ信号を用いる受信装置、受信復号化装置又は音声信号復号化装置に適用することができる。また、本発明は、移動局装置又は基地局装置にも適用することができる。

【0245】

また、本発明は上記実施の形態に限定されず、種々変更して実施することが可能である。例えば、上記実施の形態では、信号処理装置として行う場合について説明しているが、これに限られるものではなく、この信号処理方法をソフトウェアとして行うことも可能である。

【0246】

例えば、上記信号処理方法を実行するプログラムを予めROM(Read Only Memory)に格納しておき、そのプログラムをCPU(Central Processor Unit)に

よって動作させるようにしても良い。

【0247】

また、上記信号処理方法を実行するプログラムをコンピュータで読み取り可能な記憶媒体に格納し、記憶媒体に格納されたプログラムをコンピュータのRAM (Random Access memory) に記録して、コンピュータをそのプログラムにしたがって動作させるようにしても良い。

【0248】

なお、上記説明では、時間領域から周波数領域への変換法にMDCTを用いる場合について説明を行っているがこれに限定されず直交変換であればいずれも適用できる。例えば、離散フーリエ変換または離散コサイン変換等を適用することもできる。

【0249】

なお、本発明は、オーディオ信号を用いる受信装置、受信復号化装置又は音声信号復号化装置に適用することができる。また、本発明は、移動局装置又は基地局装置にも適用することができる。

【0250】

【発明の効果】

以上説明したように、本発明の音響符号化装置及び音響符号化方法によれば、入力信号をダウンサンプリングして符号化し、符号化した信号を復号化してアップサンプリングし、このアップサンプリングした復号信号と入力信号との差分信号を符号化する符号化方法において、このアップサンプリングした復号信号から拡張レイヤで符号化の対象となる周波数を決定するため、符号化側と復号化側の両方にある信号のみで前記周波数を決定することができ、よって符号化側から復号化側にこの周波数の情報を伝送ことなく、低ビットレートで高品質に符号化を行うことができる。

【図面の簡単な説明】

【図1】

本発明の実施の形態1に係る音響符号化装置の構成を示すブロック図

【図2】

音響信号の情報の分布の一例を示す図

【図3】

基本レイヤと拡張レイヤで符号化の対象とする領域の一例を示す図

【図4】

上記実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブロック図

【図5】

上記実施の形態の音響符号化装置の聴覚マスキング算出器の内部構成の一例を示す図

【図6】

上記実施の形態の拡張レイヤ符号化器の内部構成の一例を示すブロック図

【図7】

上記実施の形態の聴覚マスキング算出器の内部構成の一例を示すブロック図

【図8】

本発明の実施の形態1に係る音響復号化装置の構成を示すブロック図

【図9】

上記実施の形態の音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図

【図10】

本発明の実施の形態2の基本レイヤ符号化器の内部構成の一例を示すブロック図

【図11】

上記実施の形態の基本レイヤ復号化器の内部構成の一例を示すブロック図

【図12】

上記実施の形態の基本レイヤ復号化器の内部構成の一例を示すブロック図

【図13】

本発明の実施の形態3に係る音響符号化装置の周波数決定部の内部構成の一例を示すブロック図

【図14】

上記実施の形態の推定誤差スペクトル算出器が算出する残差スペクトルの一例を示す図

【図15】

本発明の実施の形態4に係る音響符号化装置の周波数決定部の内部構成の一例を示すブロック図

【図16】

上記実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブロック図

【図17】

本発明の実施の形態5に係る音響符号化装置の拡張レイヤ符号化器の内部構成の一例を示すブロック図

【図18】

上記実施の形態の順序づけ部の推定歪値の順位づけの一例を示す図

【図19】

本発明の実施の形態5に係る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図

【図20】

本発明の実施の形態6に係る音響符号化装置の拡張レイヤ符号化器の内部構成の一例を示すブロック図

【図21】

本発明の実施の形態6に係る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図

【図22】

上記実施の形態の音響符号化装置の周波数決定部の内部構成の一例を示すブロック図

【図23】

本発明の実施の形態6に係る音響復号化装置の拡張レイヤ復号化器の内部構成の一例を示すブロック図

【図24】

本発明の実施の形態7に係る通信装置の構成を示すブロック図

【図25】

本発明の実施の形態8に係る通信装置の構成を示すブロック図

【図26】

本発明の実施の形態9に係る通信装置の構成を示すブロック図

【図27】

本発明の実施の形態10に係る通信装置の構成を示すブロック図

【図28】

音響（音楽）信号のスペクトルの一例を示す図

【符号の説明】

101 ダウンサンプリング器

102 基本レイヤ符号化器

103 局所復号化器

104 アップサンプリング器

105 遅延器

106 減算器

107、804 周波数決定部

108 拡張レイヤ符号化器

109 多重化器

401 FFT部

402 推定聴覚マスキング算出器

403 決定部

601、701 MDCT部

602 MDCT係数量子化器

801、1101 分離器

802 基本レイヤ復号化器

803 アップサンプリング器

805 拡張レイヤ復号化器

806、903 加算器

901 MDCT係数復号化器

902 IMDCT部

1102 音源生成器

1103 合成フィルタ

1301 推定誤差スペクトル算出器

1302、1502、1601 決定部

1501 推定聴覚マスキング修正部

1701、1901 順序づけ部

1702、2002、2201 MDCT係数量子化器

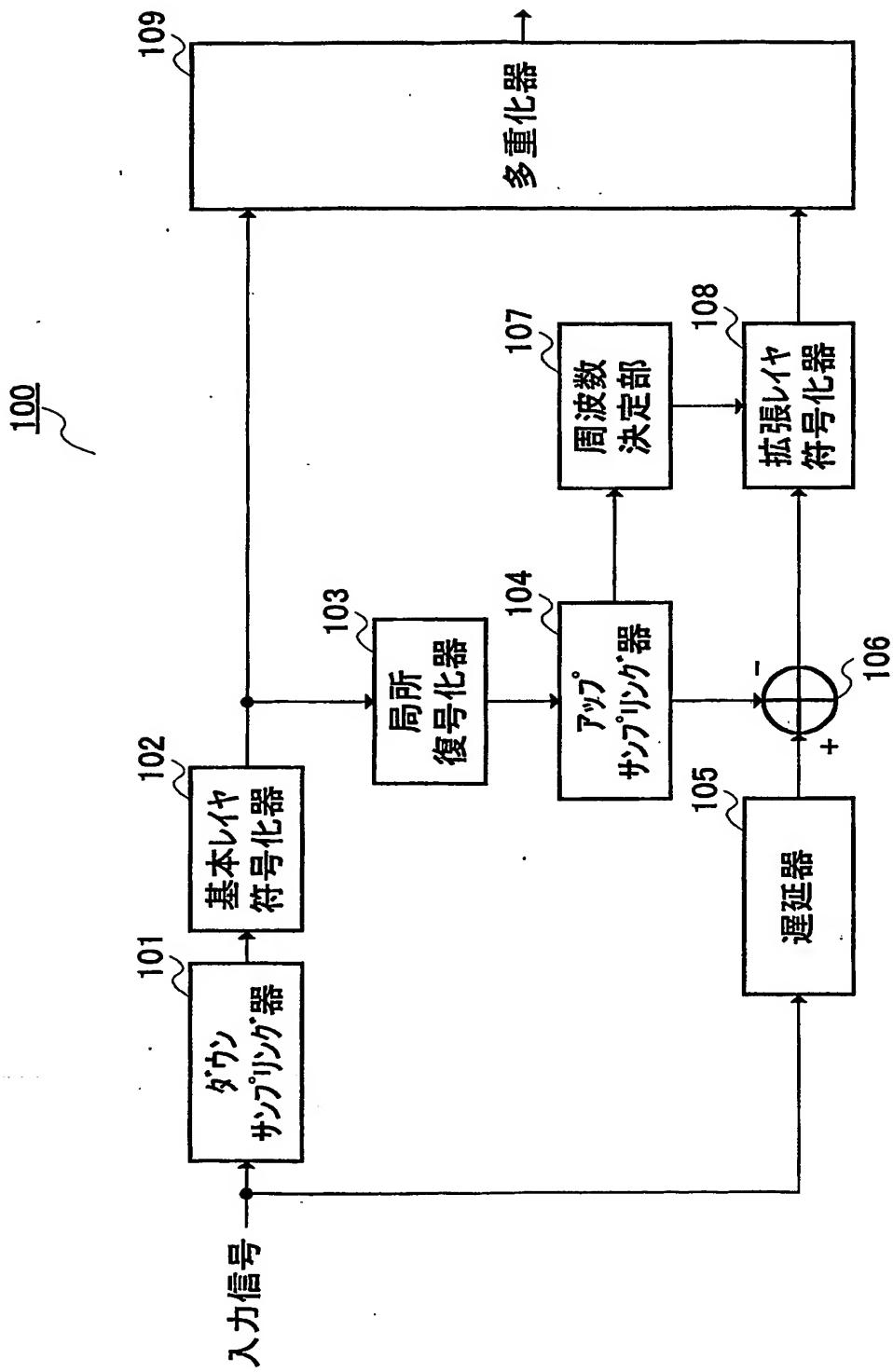
1902、2102、2301 MDCT係数復号化器

2001、2101 固定帯域指定部

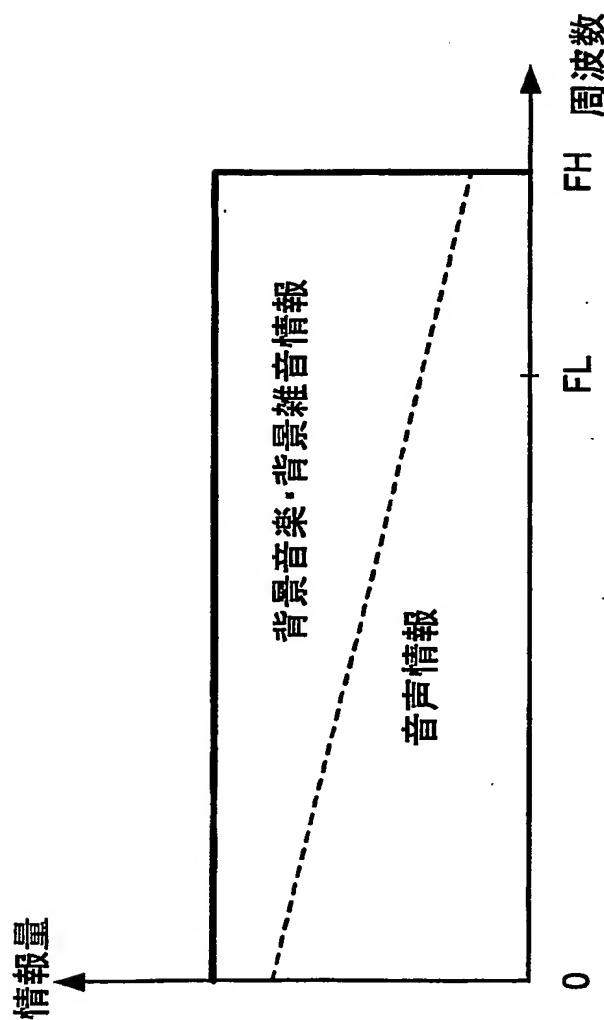
【書類名】

図面

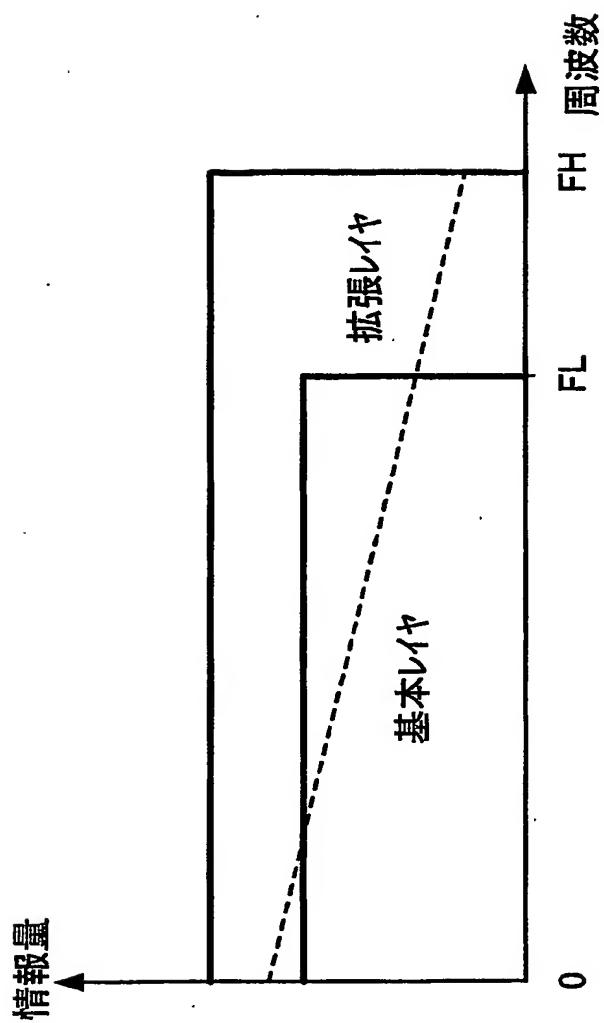
【図1】



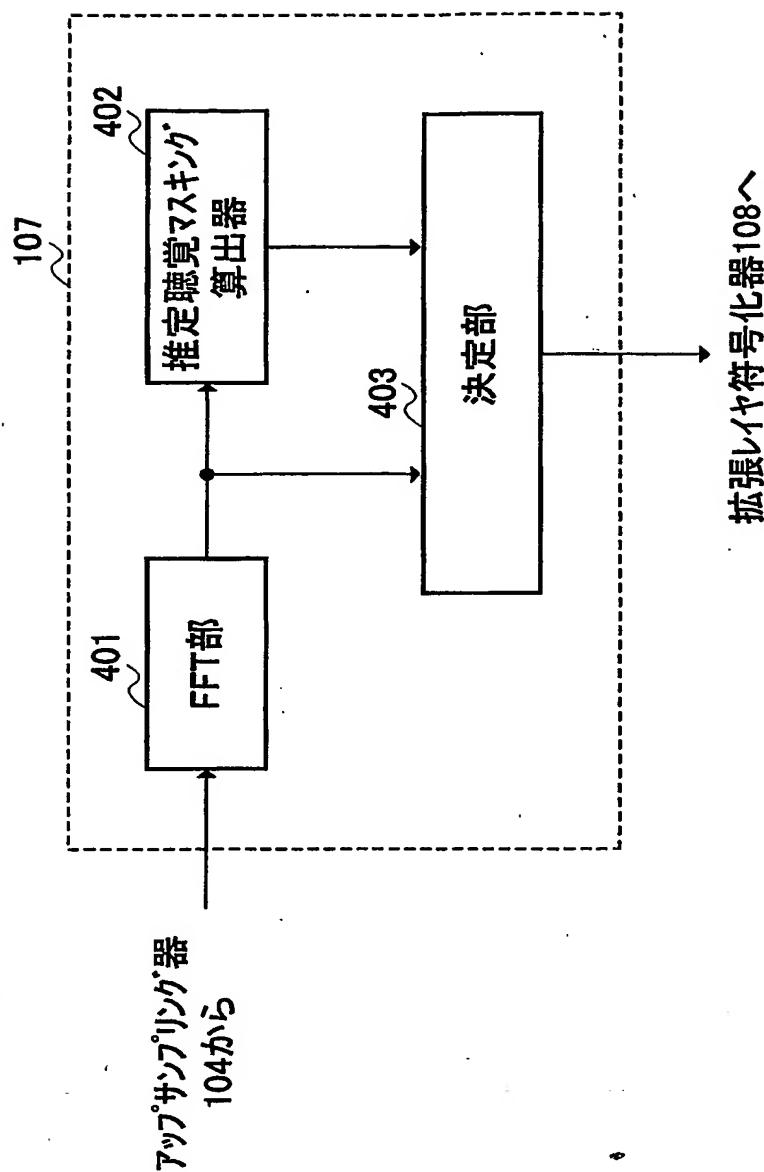
【図2】



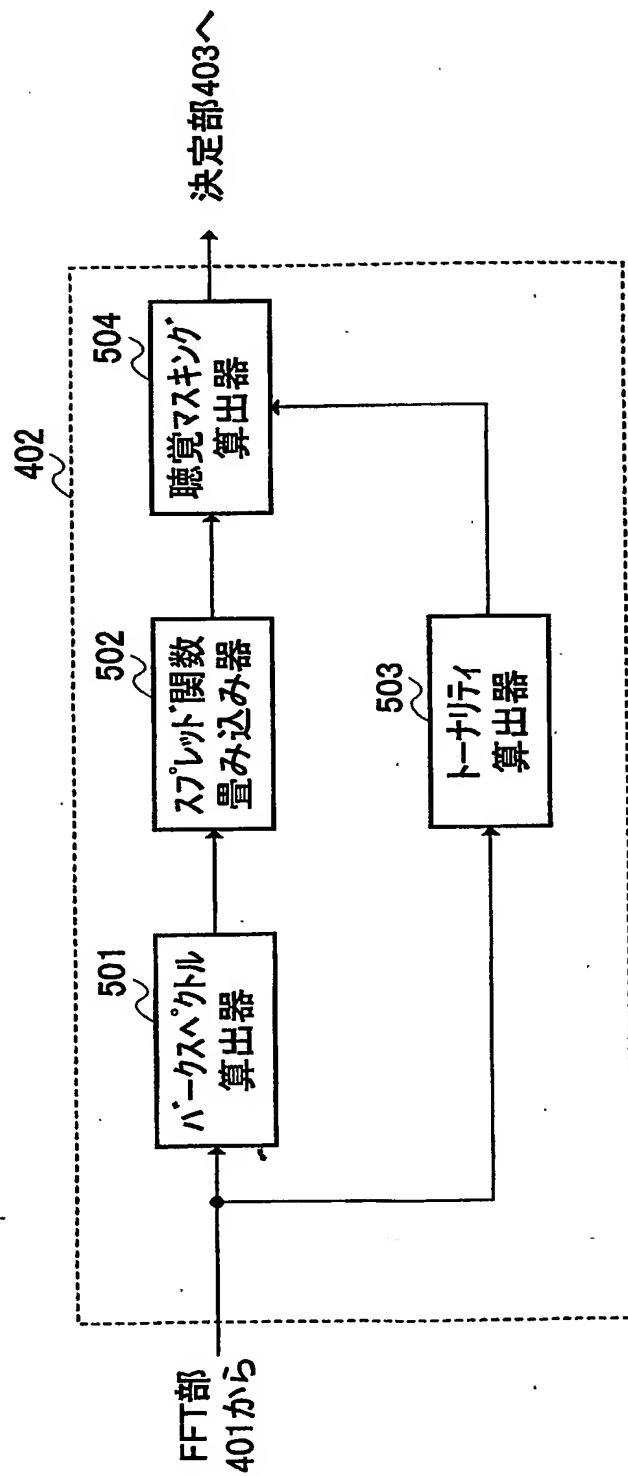
【図3】



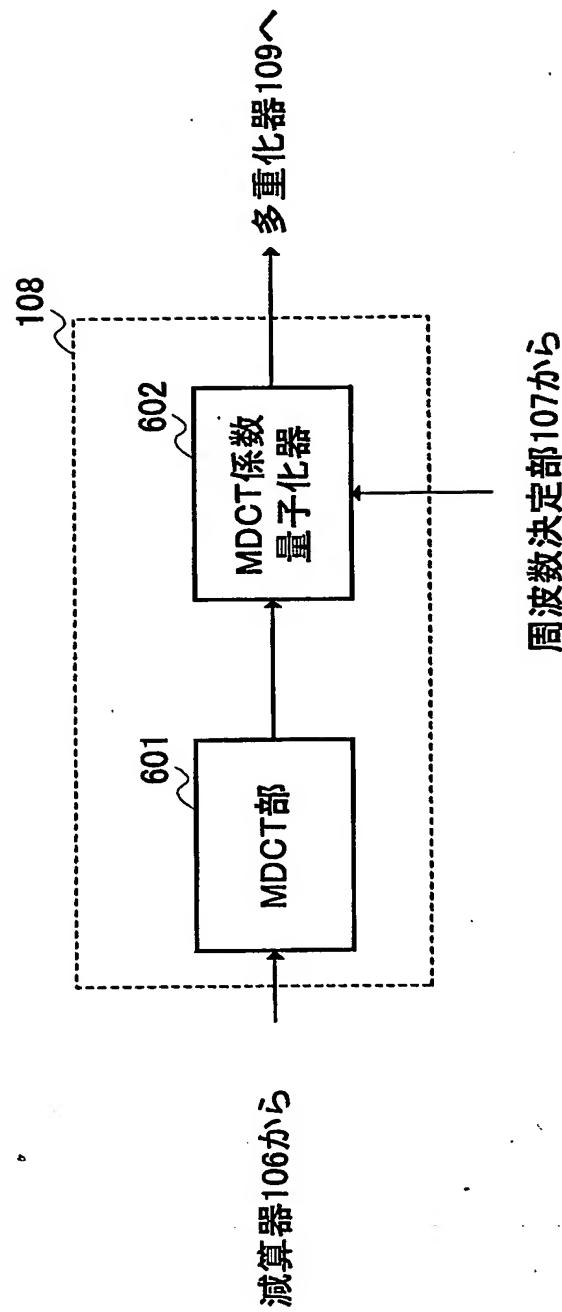
【図4】



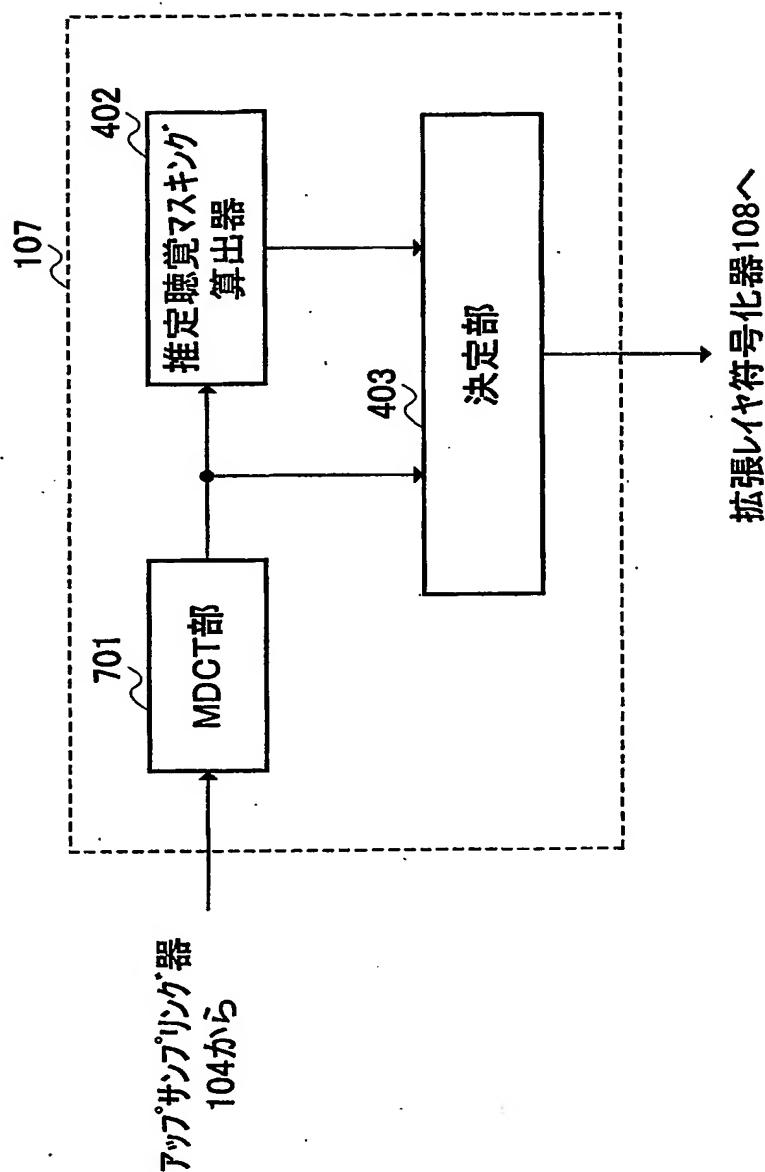
【図5】



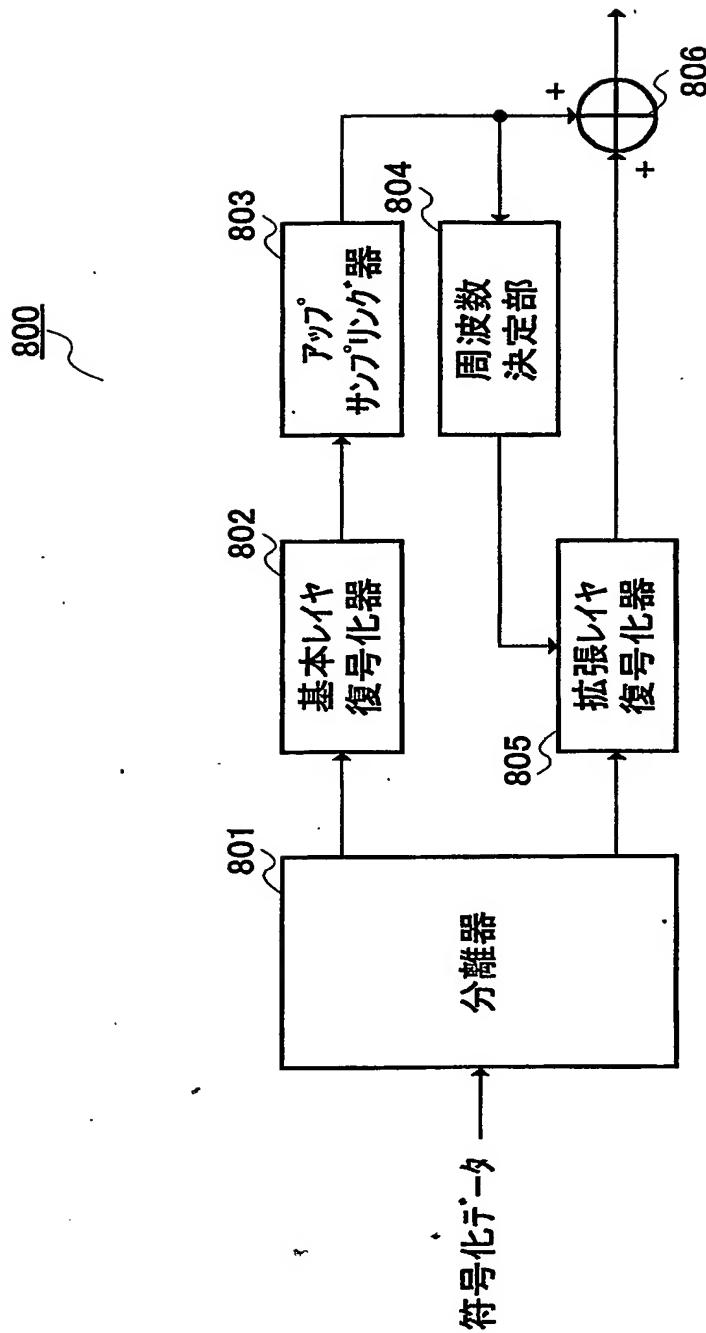
【図6】



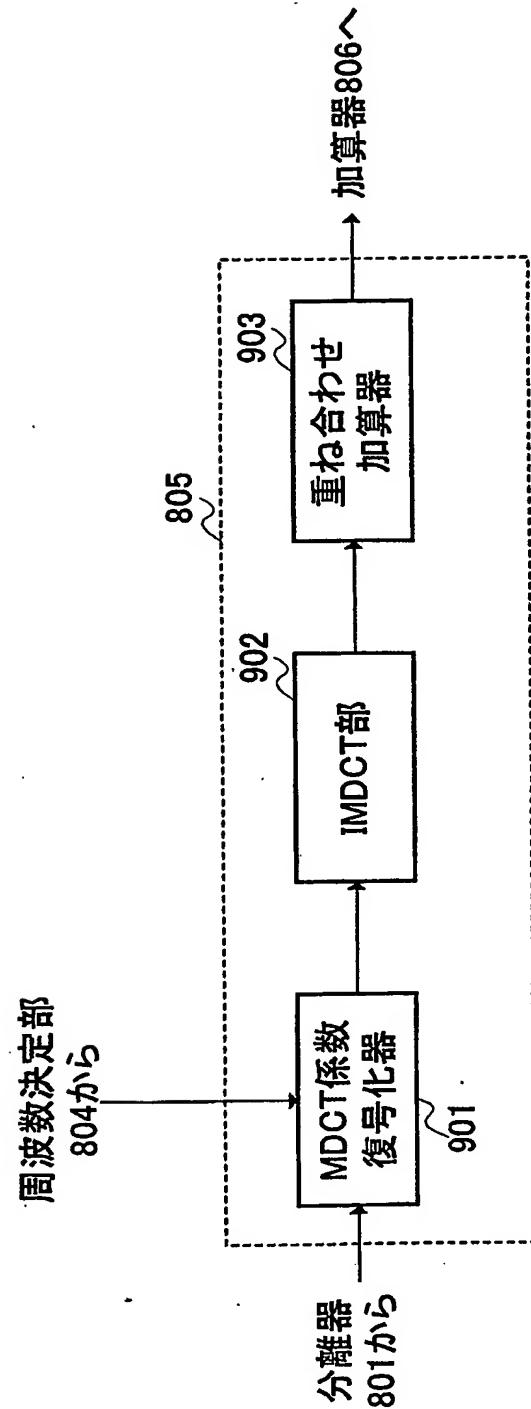
【図7】



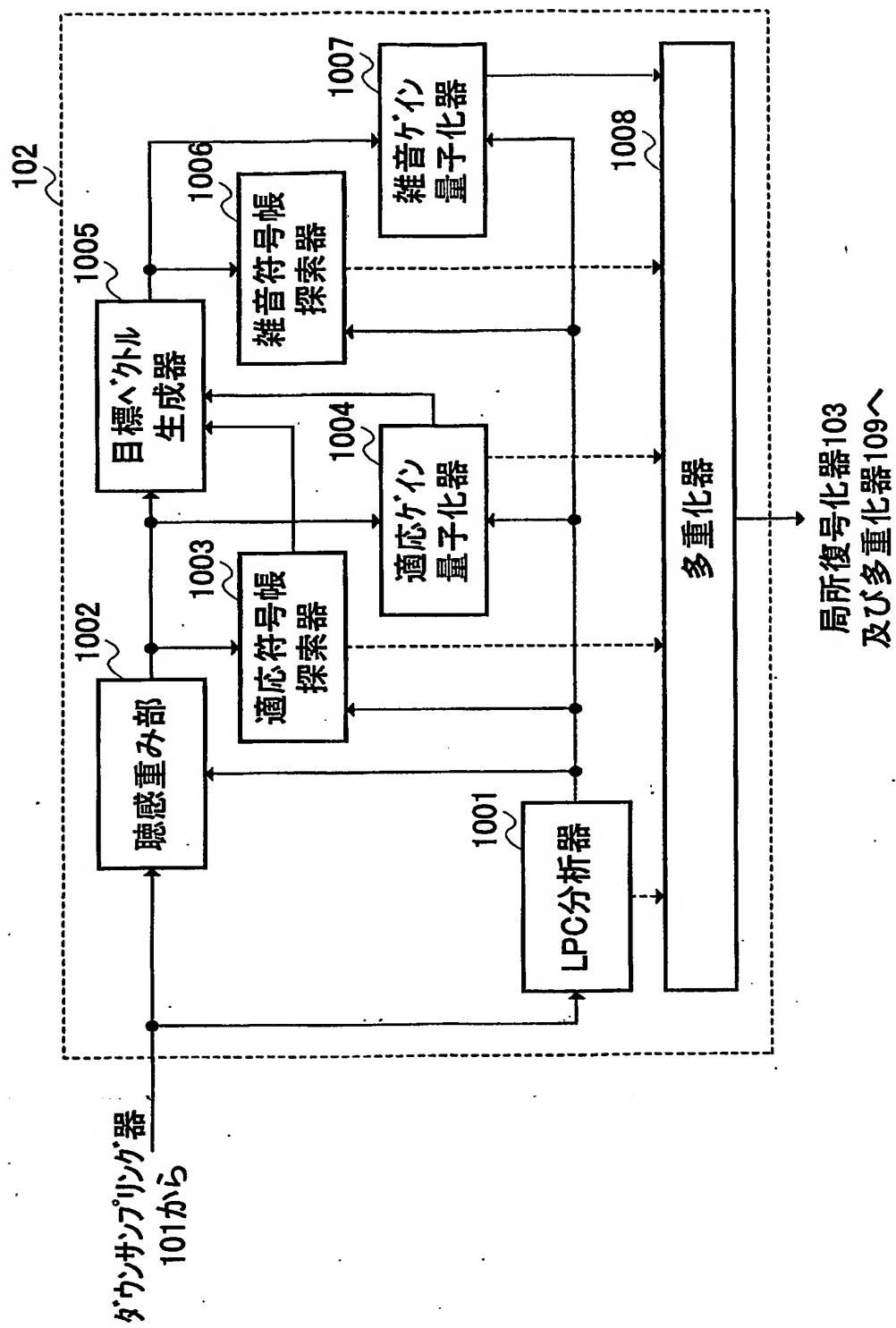
【図8】



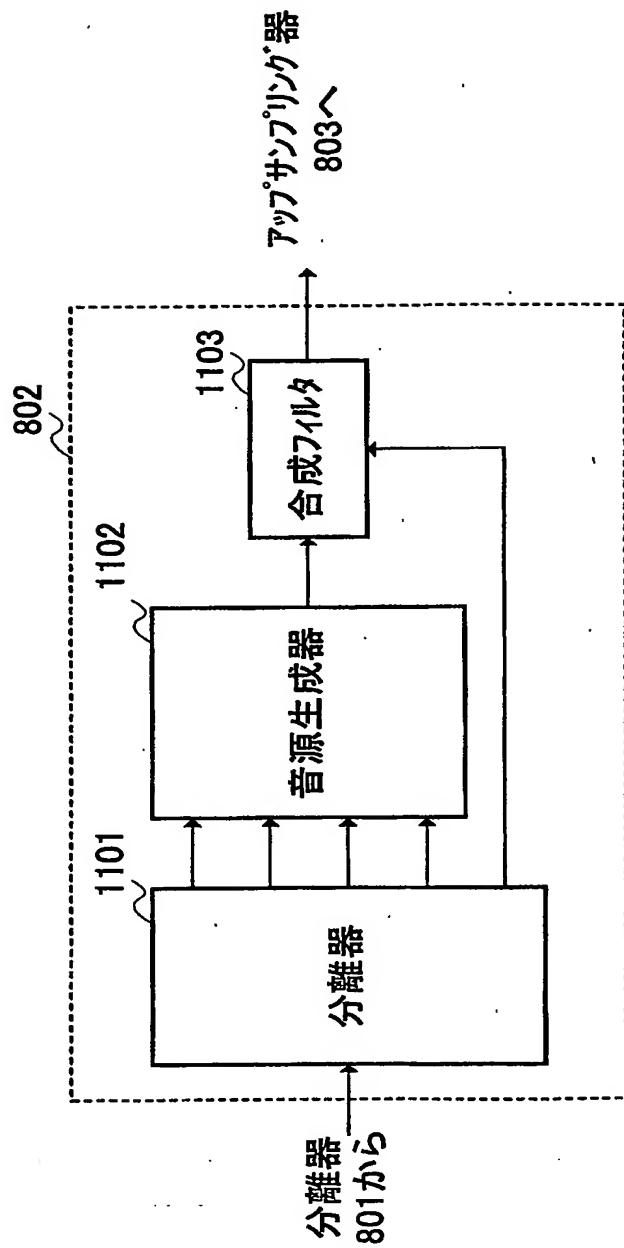
【図9】



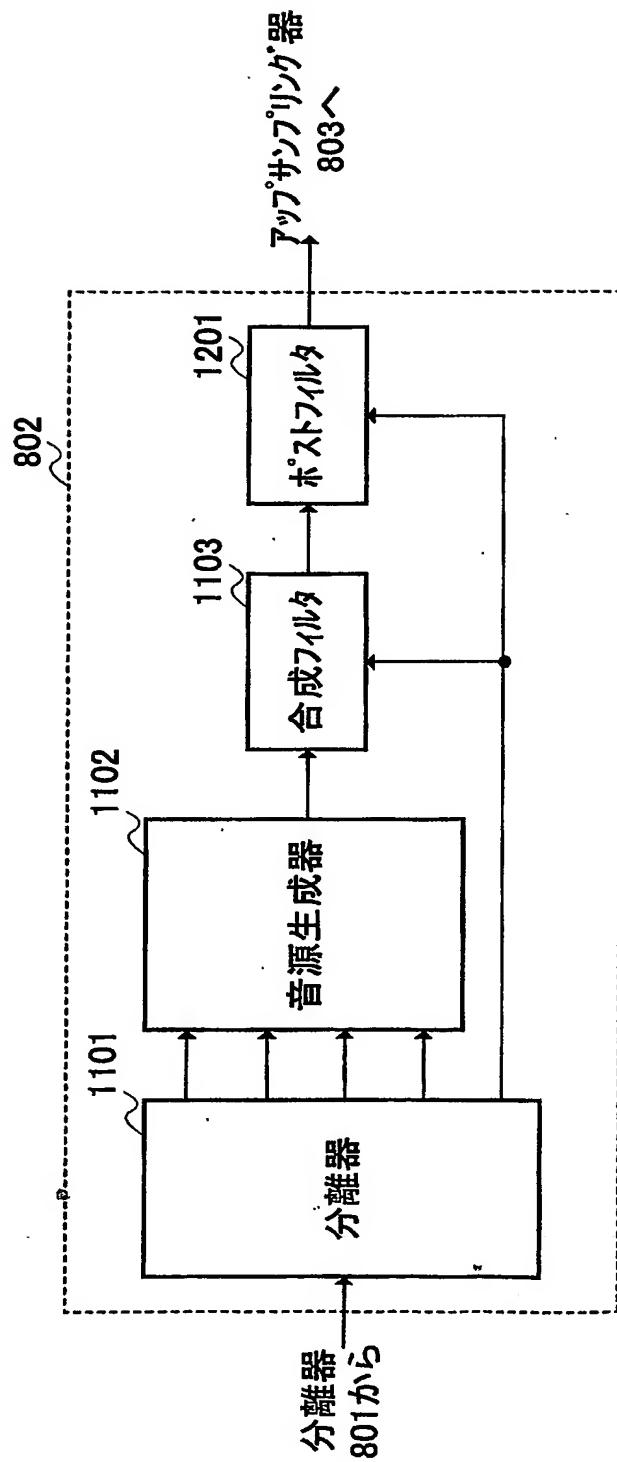
【図10】



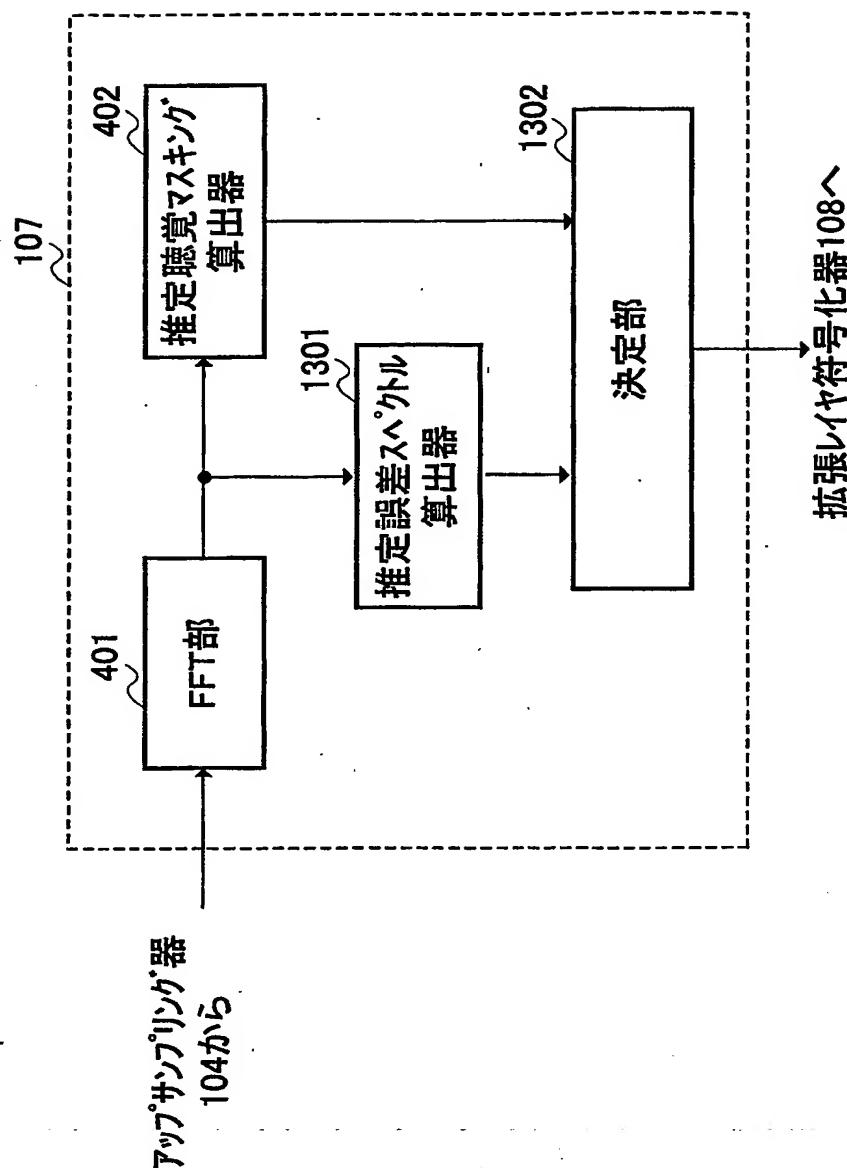
【図11】



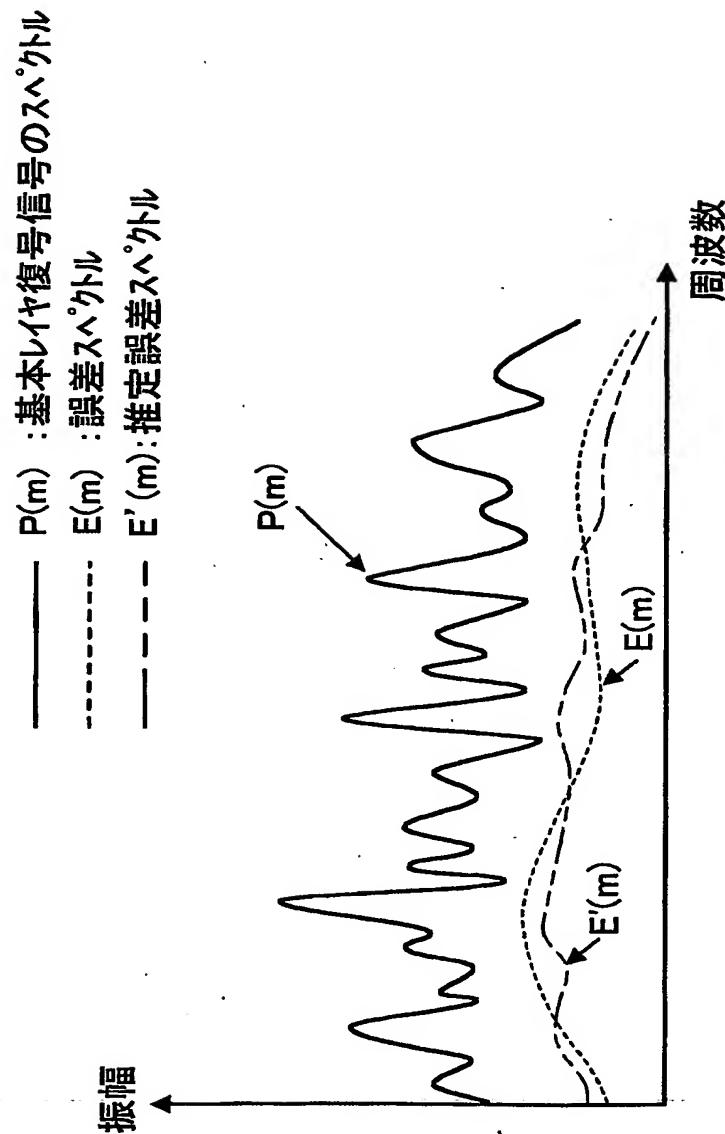
【図12】



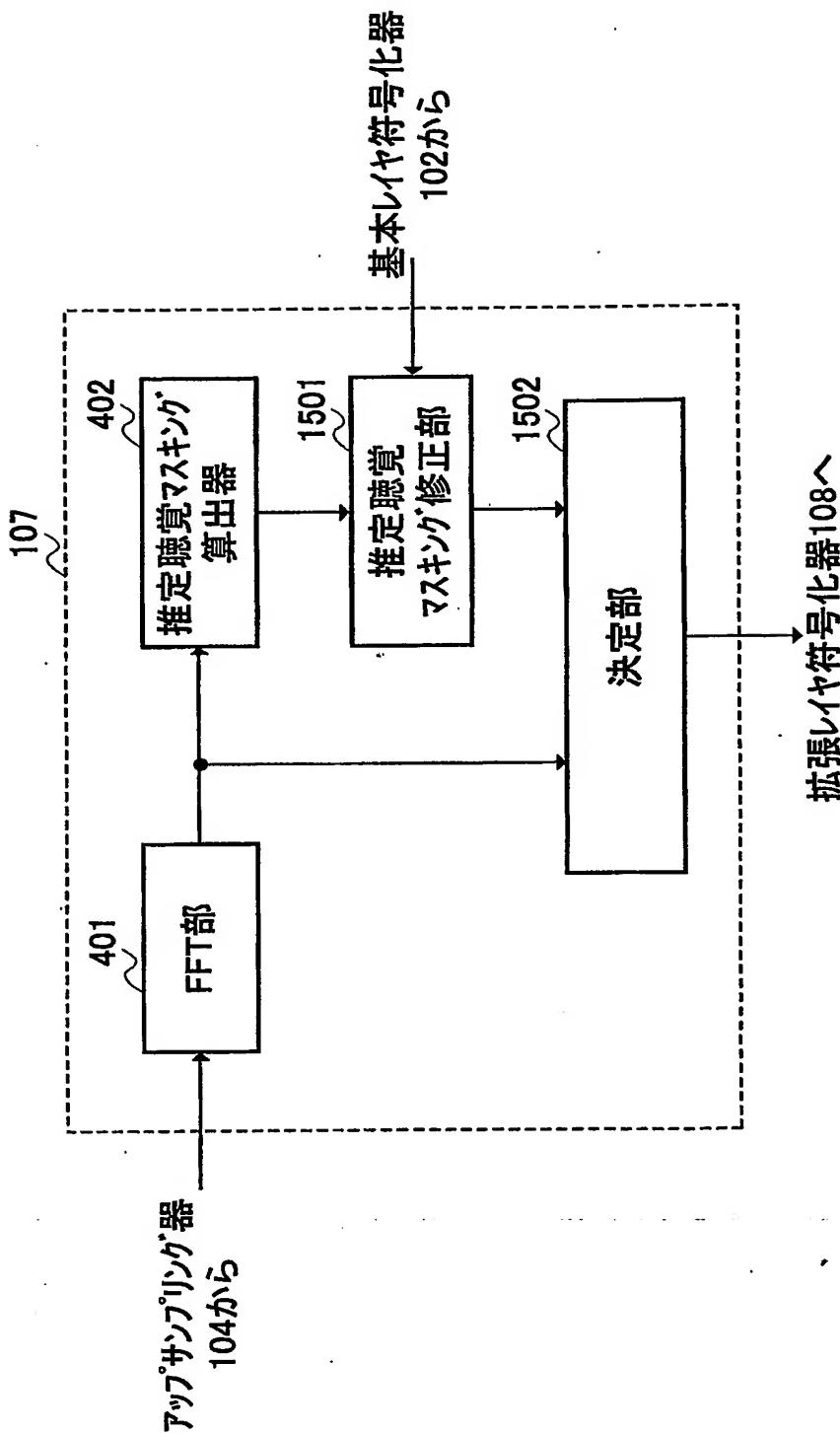
【図13】



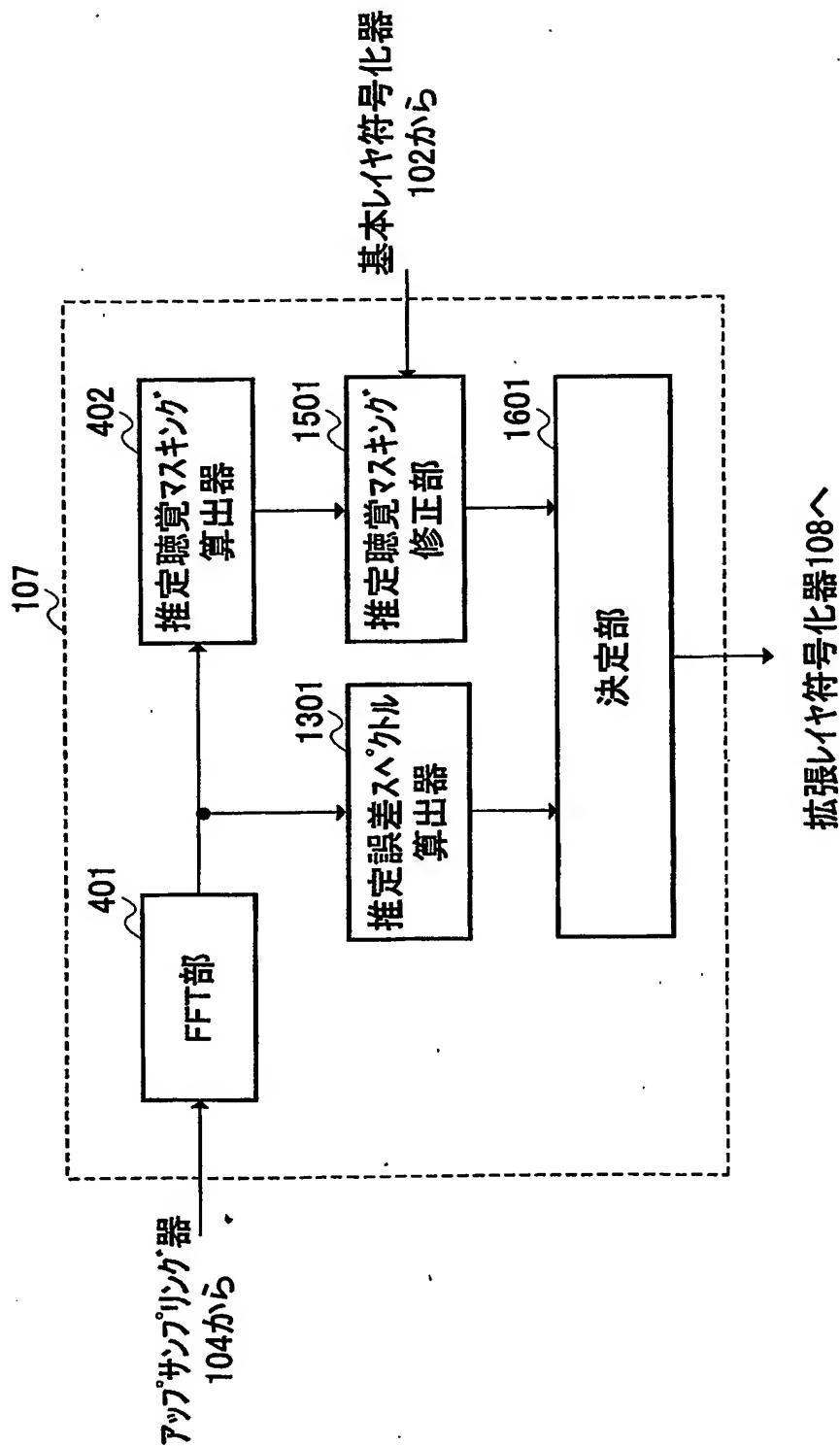
【図14】



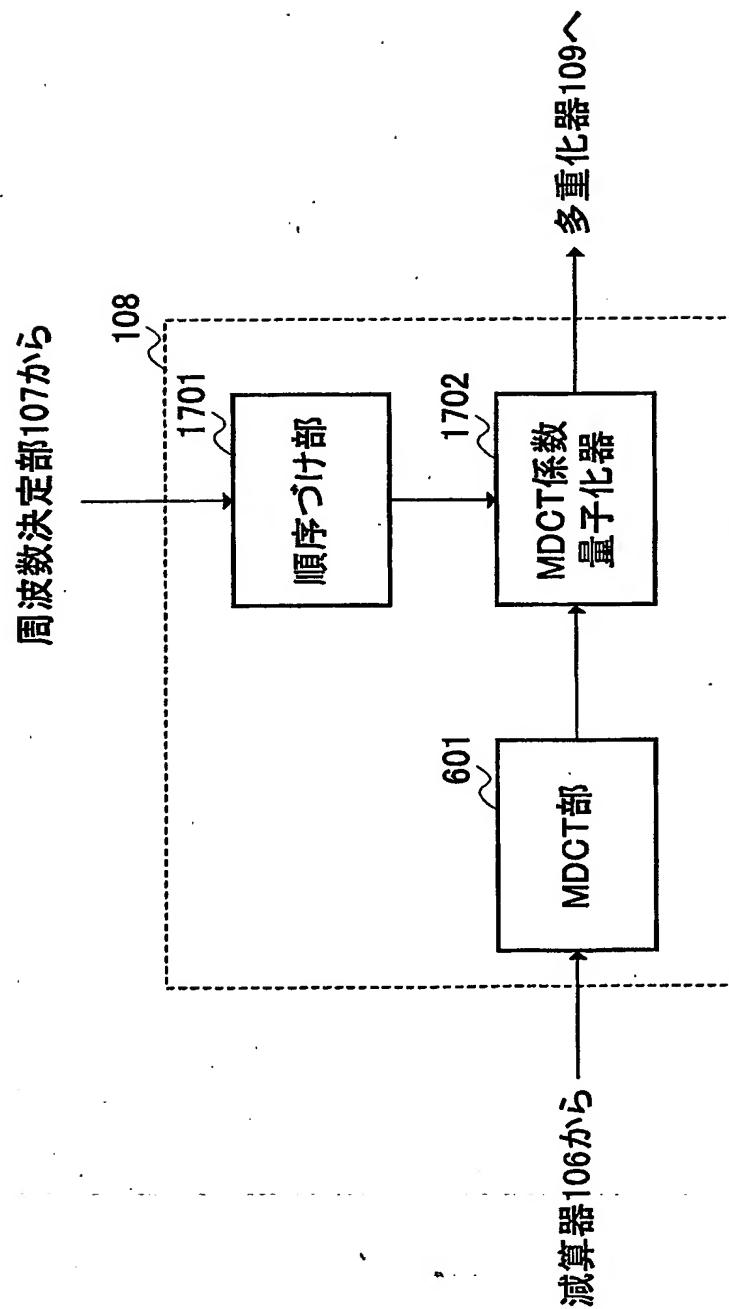
【図15】



【図16】



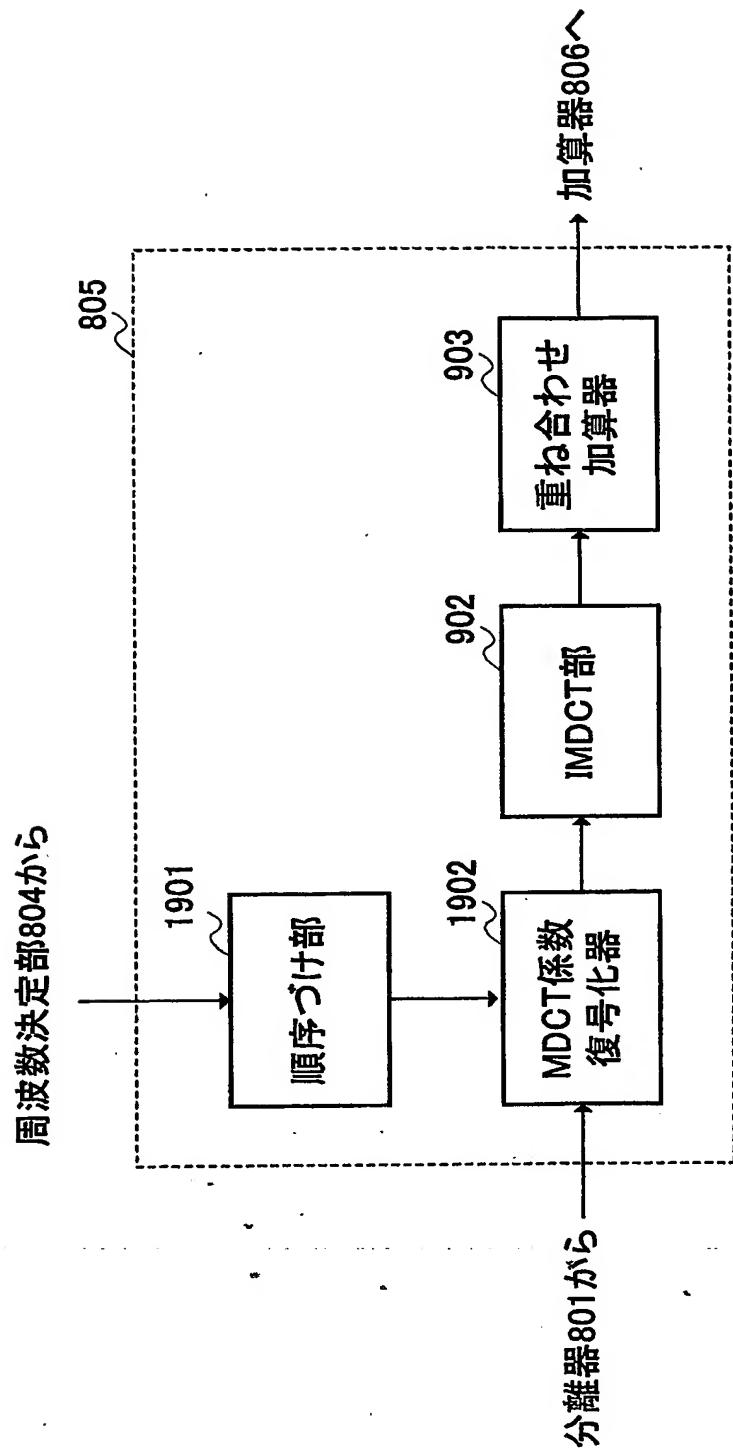
【図17】



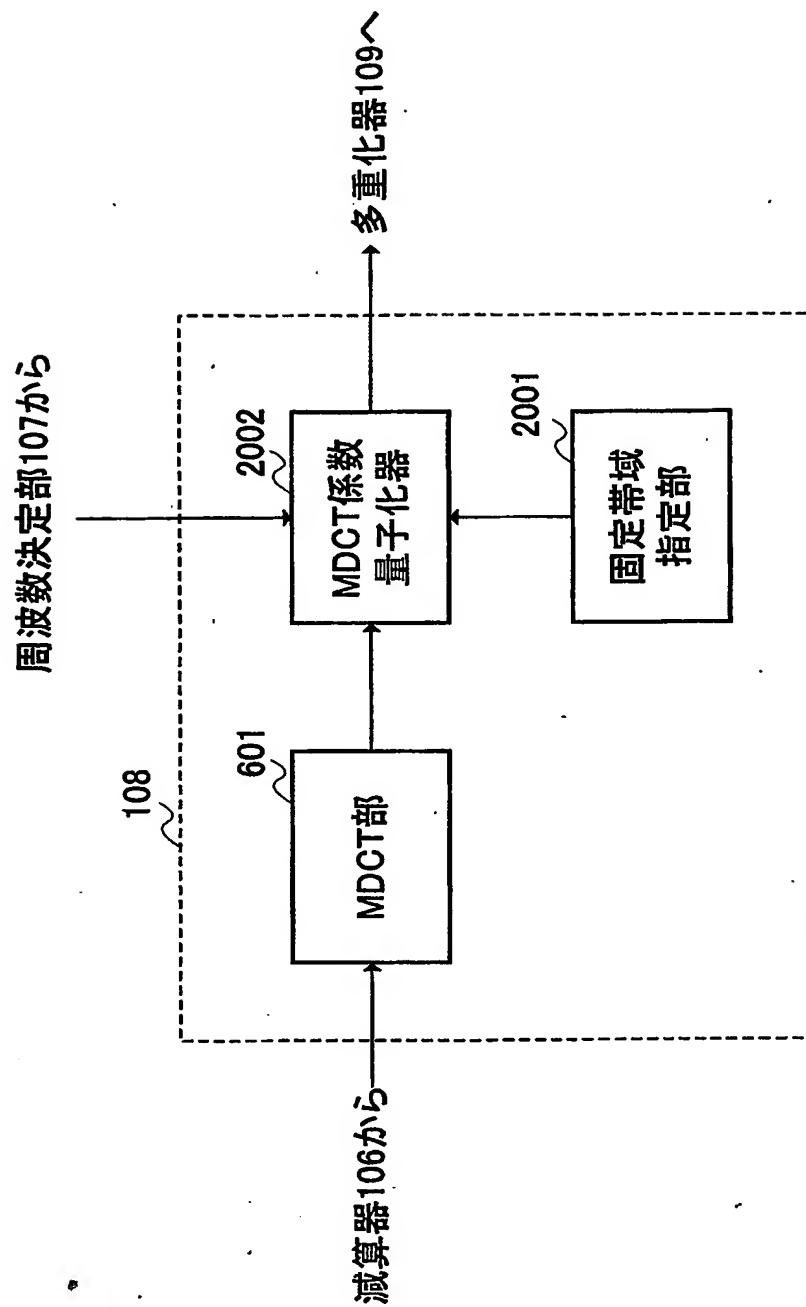
【図18】

周波数 (m)	推定歪値 D(m)	順序
1	24.0	5
3	16.8	7
4	35.9	3
7	147.2	1
8	135.8	2
9	26.9	4
11	23.0	6
12	12.9	8

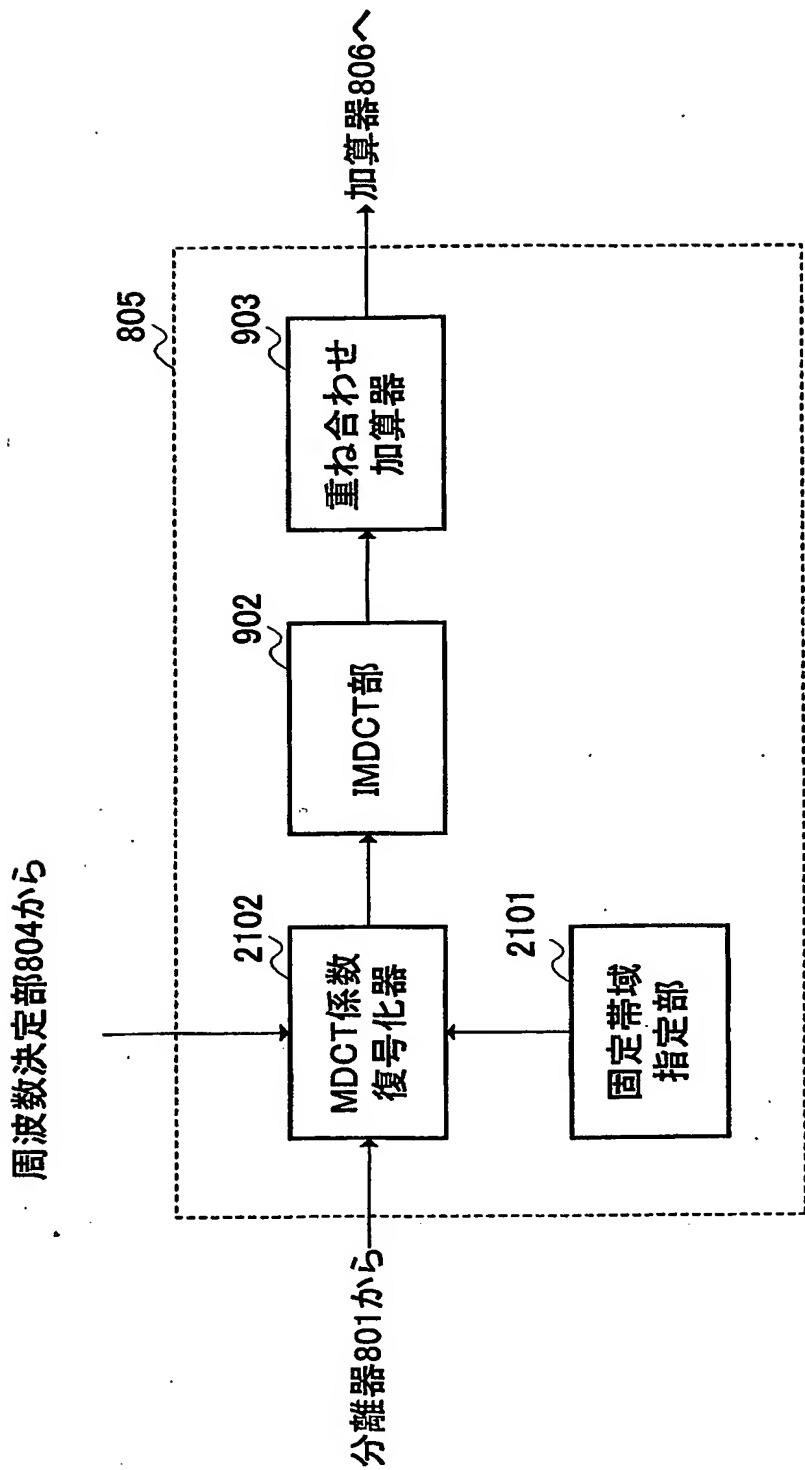
【図19】



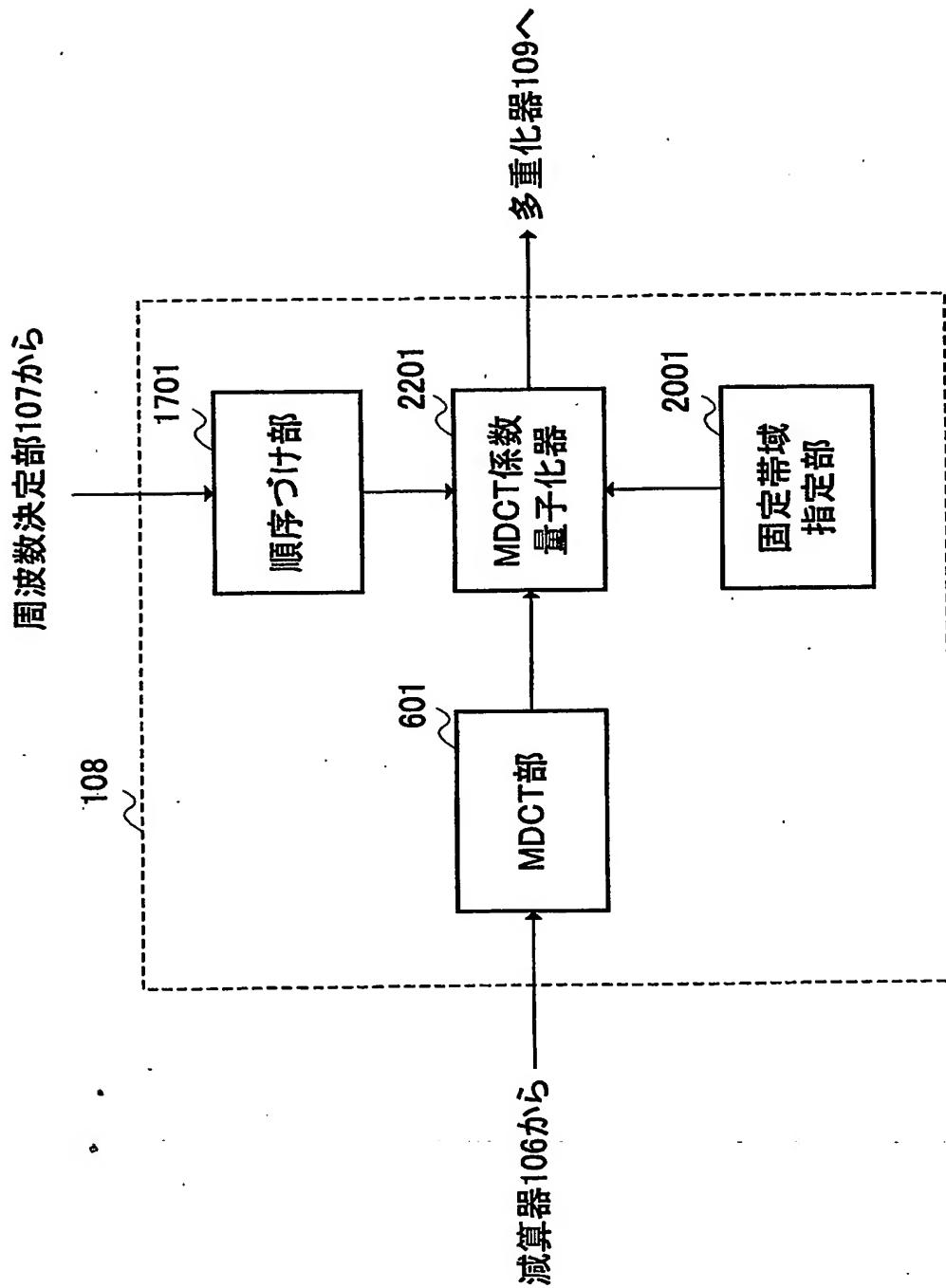
【図20】



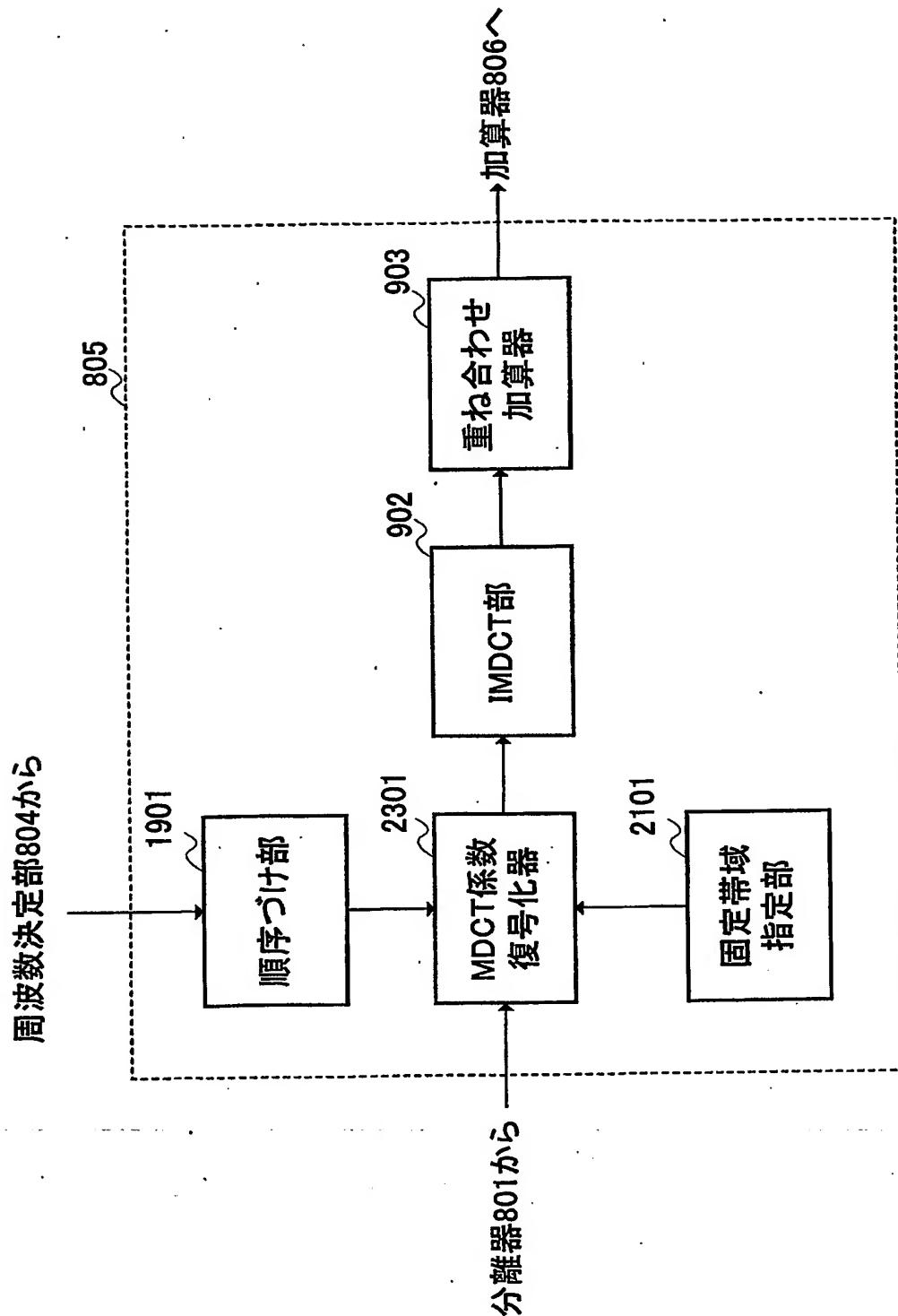
【図21】



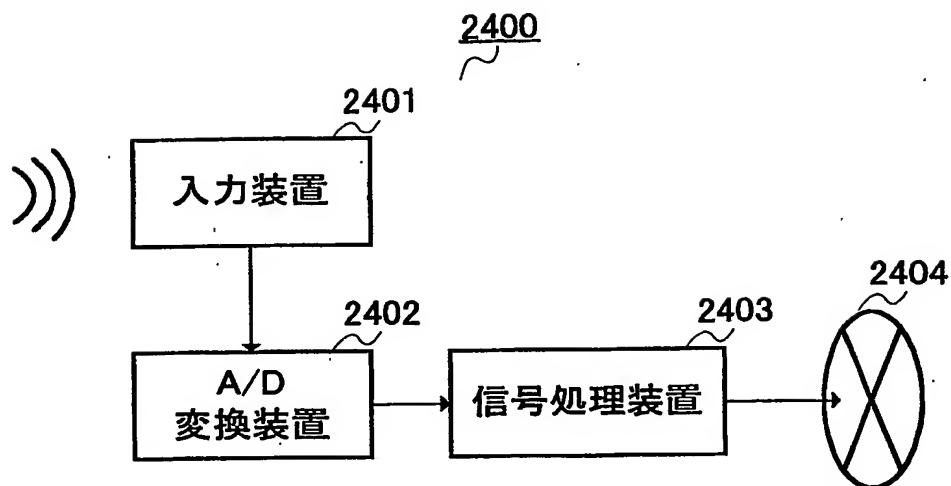
【図22】



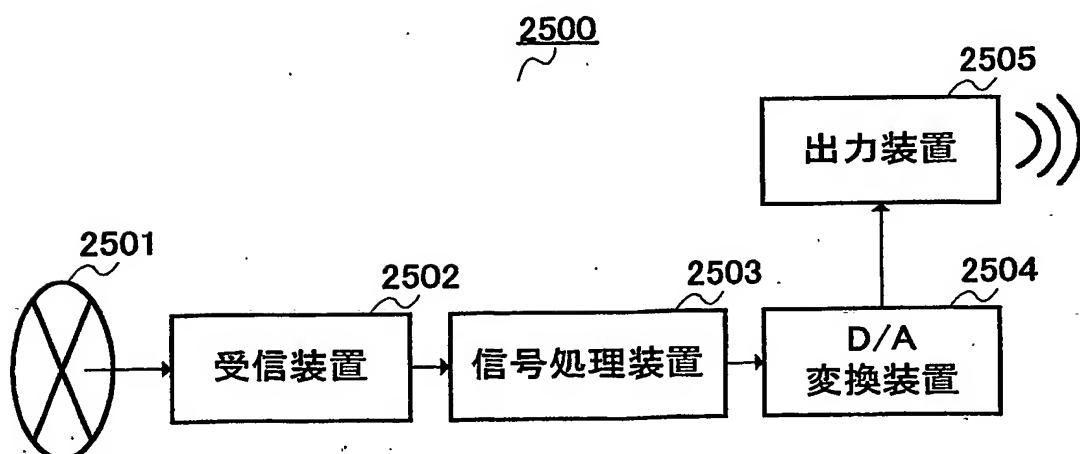
【図23】



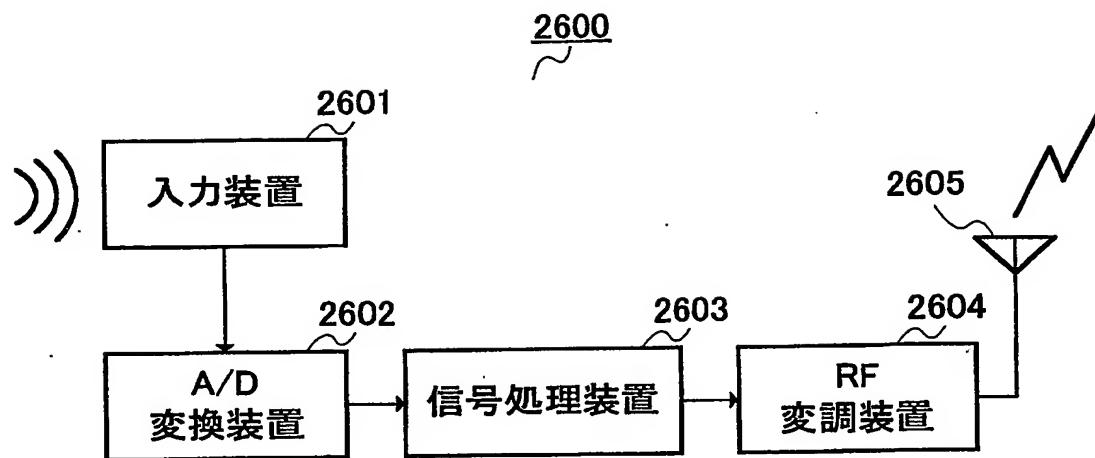
【図24】



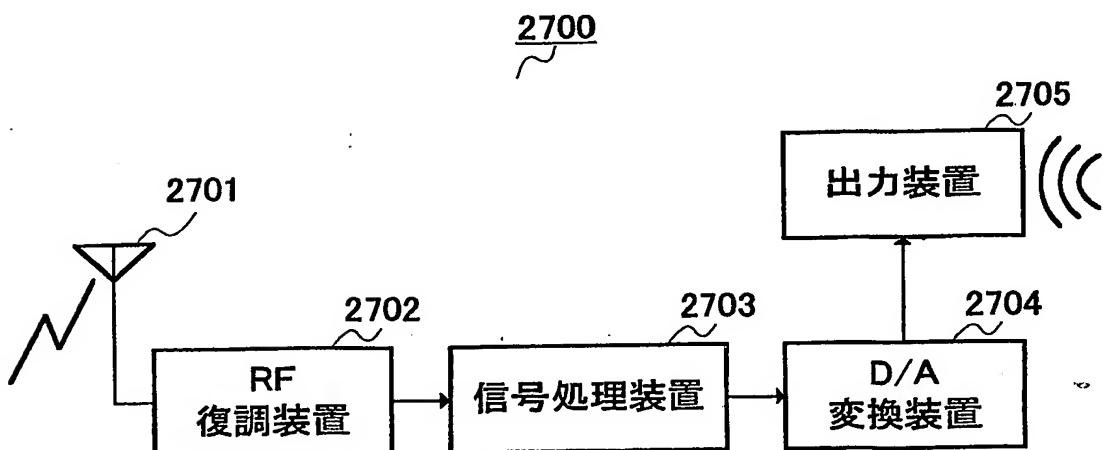
【図25】



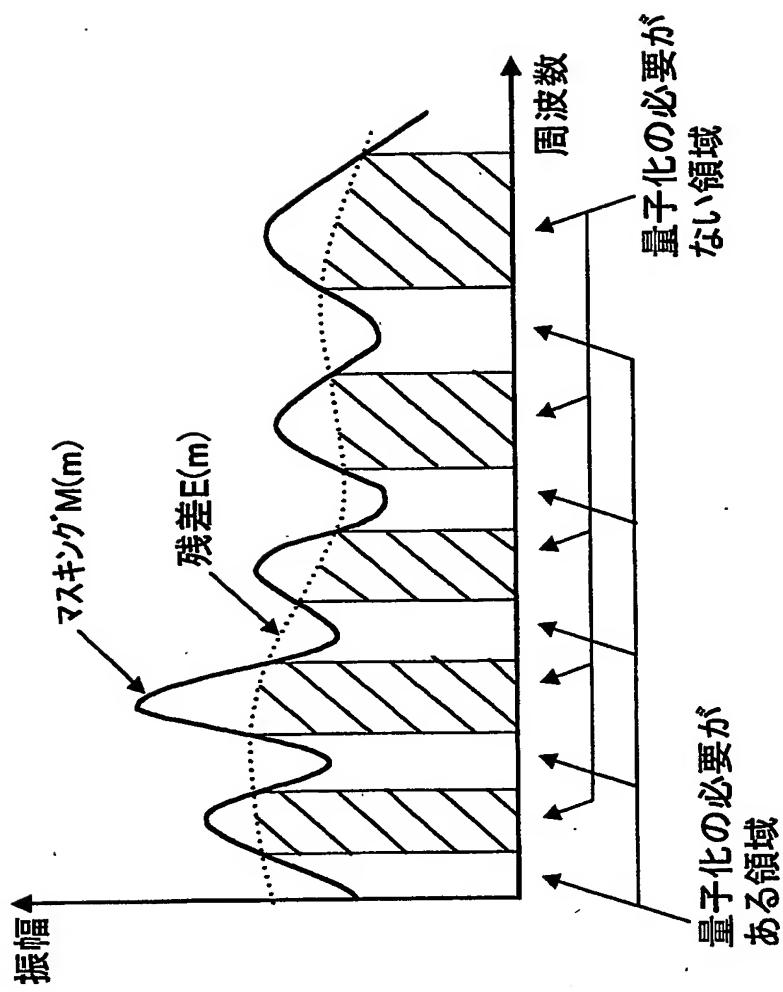
【図26】



【図27】



【図28】



【書類名】 要約書

【要約】

【課題】 声が主体で背景に音楽や雑音が重畠しているような信号であっても、低ビットレートで高品質に符号化を行うこと。

【解決手段】 基本レイヤ符号化器102は、サンプリングレートFLの入力データを所定の基本フレーム単位で符号化して第1符号化コードを生成する。局所復号化器103は、第1符号化コードを復号化する。アップサンプリング器104は、復号信号のサンプリングレートをFHに上げる。周波数決定部107は、サンプリングレートをFHに上げた復号信号から誤差信号の符号化する領域と、符号化しない領域を決定する。拡張レイヤ符号化器108は、誤差信号を周波数領域の係数に変換して誤差スペクトルを生成し、周波数決定部107から得られる符号化の対象となる周波数情報に基づき誤差スペクトルの符号化を行う。

【選択図】 図1

出願人履歴情報

識別番号 [000005821]

1. 変更年月日 1990年 8月28日

[変更理由] 新規登録

住 所 大阪府門真市大字門真1006番地

氏 名 松下電器産業株式会社

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.